Epistasis, inbreeding depression and the evolution of self-fertilization

Diala Abu Awad* and Denis Roze†,‡

* Department of Population Genetics, Technical University of Munich, Germany

† CNRS, UMI 3614 Evolutionary Biology and Ecology of Algae, 29688 Roscoff, France

‡ Sorbonne Université, Station Biologique de Roscoff, 29688 Roscoff, France

Running title: Epistasis and selfing evolution

Keywords: epistasis, evolutionary quantitative genetics, inbreeding depression, multilocus population genetics, pollen discounting, self-fertilization

Address for correspondence:

Denis Roze

Station Biologique de Roscoff

Place Georges Teissier, CS90074

29688 Roscoff Cedex

France

Phone: (+33) 2 56 45 21 39

Fax: (+33) 2 98 29 23 24

email: roze@sb-roscoff.fr

ABSTRACT

Inbreeding depression resulting from partially recessive deleterious alleles is thought to be the main genetic factor preventing self-fertilizing mutants from spreading in outcrossing hermaphroditic populations. However, deleterious alleles may also generate an advantage to selfers in terms of more efficient purging, while the effects of epistasis among those alleles on inbreeding depression and mating system evolution remain little explored. In this paper, we use a general model of selection to disentangle the effects of different forms of epistasis (additive-by-additive, additive-by-dominance and dominance-by-dominance) on inbreeding depression and on the strength of selection for selfing. Models with fixed epistasis across loci, and models of stabilizing selection acting on quantitative traits (generating distributions of epistasis) are considered as special cases. Besides its effects on inbreeding depression, epistasis may increase the purging advantage associated with selfing (when it is negative on average), while the variance in epistasis favors selfing through the generation of linkage disequilibria that increase mean fitness. Approximations for the strengths of these effects are derived, and compared with individual-based simulation results.

INTRODUCTION

Self-fertilization is a widespread mating system found in hermaphroditic plants and animals (e.g., Jarne and Auld, 2006; Igic and Busch, 2013). In Angiosperms, the transition from outcrossing to selfing occurred multiple times, leading to approximately $10-15\%$ of species self-fertilizing at very high rates (Barrett et al., 2014). Two possible benefits of selfing have been proposed to explain such transitions: the possibility for a single individual to generate offspring in the absence of mating partner or pollinator ("reproductive assurance", Darwin, 1876; Stebbins, 1957; Porcher and Lande, 2005a; Busch and Delph, 2012), and the "automatic advantage" stemming from the fact that, in a population containing both selfers and outcrossers, selfers tend to transmit more copies of their genome to the next generation if they continue to export pollen — thus retaining the ability to sire outcrossed ovules (Fisher, 1941; Charlesworth, 1980; Stone et al., 2014). The main evolutionary force thought to oppose the spread of selfing is inbreeding depression, the decreased fitness of inbred offspring resulting from the expression of partially recessive deleterious alleles segregating within populations (Charlesworth and Charlesworth, 1987). When selfers export as much pollen as outcrossers (leading to a 50% transmission advantage for selfing), inbreeding depression must be 0.5 to compensate for the automatic advantage of selfing (Lande and Schemske, 1985). However, observations from natural populations indicate that self-fertilizing individuals do not always export as much pollen as their outcrossing counterparts, as some of their pollen production is used to fertilize their own ovules (see references in Porcher and Lande, 2005a). This phenomenon, known as pollen discounting, decreases the automatic advantage of selfing (Nagylaki, 1976; Charlesworth, 1980), thus

reducing the threshold value of inbreeding depression above which outcrossing can be maintained (e.g., Holsinger et al., 1984). It may also lead to evolutionarily stable mixed mating systems (involving both selfing and outcrossing) under some models of discounting such as the mass-action pollination model (Holsinger, 1991; Porcher and Lande, 2005a).

Several models explored the evolution of mating systems while explicitly representing the genetic architecture of inbreeding depression (e.g., Charlesworth et al., 1990; Uyenoyama and Waller, 1991; Epinat and Lenormand, 2009; Porcher and Lande, 2005b; Gervais et al., 2014), and highlighted the importance of another genetic factor (besides the automatic advantage and inbreeding depression) affecting the evolution of selfing. This third factor stems from the fact that selection against deleterious alleles is more efficient among selfed offspring (due to their increased homozygosity) than among outcrossed offspring, generating positive linkage disequilibria between alleles increasing the selfing rate and the more advantageous alleles at selected loci. Alleles increasing selfing thus tend to be found on better purged genetic backgrounds, which may allow selfing to spread even when inbreeding depression is higher than 0.5 (Charlesworth et al., 1990). This effect becomes more important as the strength of selection against deleterious alleles increases (so that purging occurs more rapidly), recombination decreases, and as alleles increasing selfing have larger effects — so that linkage disequilibria can be maintained over larger numbers of generations (Charlesworth et al., 1990; Uyenoyama and Waller, 1991; Epinat and Lenormand, 2009). This corresponds to Lande and Schemske's (1985) verbal prediction that a mutant allele coding for complete selfing may increase in frequency regardless of the amount of inbreeding depression.

Most genetic models on the evolution of selfing assume that deleterious alleles

at different loci have multiplicative effects (no epistasis). Charlesworth et al. (1991) considered a deterministic model including synergistic epistasis between deleterious alleles, showing that this form of epistasis tends to flatten the relation between inbreeding depression and the population's selfing rate, inbreeding depression sometimes increasing at high selfing rates. Concerning the spread of selfing modifier alleles, the results were qualitatively similar to the multiplicative model, except that, for parameter values where full outcrossing is not stable, the evolutionarily stable selfing rate tended to be slightly below 1 under synergistic epistasis (whereas it would have been at exactly 1 in the absence of epistasis). Other models explored the effect of partial selfing on inbreeding depression generated by polygenic quantitative traits under stabilizing selection (Lande and Porcher, 2015; Abu Awad and Roze, 2018). This type of model typically generates distributions of epistatic interactions across loci, including possible compensatory effects between mutations. When effective recombination is sufficiently weak, linkage disequilibria generated by epistasis may greatly reduce inbreeding depression, and even generate outbreeding depression between selfing lineages carrying different combinations of compensatory mutations. However, the evolution of the selfing rate was not considered by these models.

In this paper, we use a general model of epistasis between pairs of selected loci to explore the effects of epistasis on inbreeding depression and on the evolution of selfing. We derive analytical approximations showing that epistatic interactions affect the spread of selfing modifiers through various mechanisms: by affecting inbreeding depression, the purging advantage of selfers and also through linkage disequilibria between selected loci. Although the expressions obtained can become complicated for intermediate selfing rates, we will see that the condition determining whether selfing can spread

in a fully outcrossing population often remains relatively simple. Notably, our model allows us to disentangle the effects of additive-by-additive, additive-by-dominance and dominance-by-dominance epistatic interactions on inbreeding depression and selection for selfing — while the models used by Charlesworth et al. (1991), Lande and Porcher (2015) and Abu Awad and Roze (2018) impose certain relations between these quantities. The cases of fixed, synergistic epistasis and of stabilizing selection acting on quantitative traits (Fisher's geometric model) will be considered as special cases, for which we will also present individual-based simulation results. Overall, our results show that, for a given level of inbreeding depression and average strength of selection against deleterious alleles, epistatic interactions tend to facilitate the spread of selfing, due to the fact that selfing can maintain beneficial combinations of alleles.

METHODS

Life cycle. Our analytical model represents an infinite, hermaphroditic population with discrete generations. A proportion σ of ovules produced by a given individual are self-fertilized, while its remaining ovules are fertilized by pollen sampled from the population pollen pool (Table 1 provides a list of the symbols used throughout the paper). A parameter κ represents the rate of pollen discounting: an individual with selfing rate σ contributes to the pollen pool in proportion $1 - \kappa\sigma$ (e.g., Charlesworth, 1980). Therefore, κ equals 0 in the absence of pollen discounting, while κ equals 1 under full discounting (in which case complete selfers do not contribute to the pollen pool). We assume that the selfing rate σ is genetically variable, and coded by ℓ_σ loci

with additive effects:

$$\sigma = \sum_{i=1}^{\ell_\sigma} \left(\sigma_i^{\mathrm{M}} + \sigma_i^{\mathrm{P}}\right) \tag{1}$$

where the sum is over all loci affecting the selfing rate, and where σ_i^{M} and σ_i^{P} represent the effect of the alleles present respectively on the maternally and paternally inherited genes at locus i (note that the assumption of additivity within and between loci may not always hold, in particular when selfing rates are close to 0 or 1). The model does not make any assumption concerning the number of alleles segregating at loci affecting the selfing rate; however, our analysis will assume that the variance of σ in the population remains small and that linkage disequilibria between loci affecting the selfing rate may be neglected, effectively leading to the same expression for the selection gradient on the selfing rate as in a simpler model considering the spread of a mutant allele changing σ by a small amount. Although we assume that the selfing rate is purely genetically determined, our general results should still hold when σ is also affected by (uncorrelated) environmental effects, after multiplying expressions for the change in the average selfing rate over time by the heritability of σ.

The fitness W of an organism is defined as its overall fecundity (that may depend on its survival), so that the expected number of seeds produced by an individual is proportional to W, while its contribution to the population pollen pool is proportional to $W\,(1-\kappa\sigma)$. We assume that W is affected by a possibly large number ℓ of biallelic loci. Alleles at each of these loci are denoted 0 and 1; we assume an equal mutation rate u from 0 to 1 and from 1 to 0, assumed to be small relative to the strength of selection at each locus. The overall mutation rate (per haploid genome) at loci affecting fitness is denoted $U = u\,\ell$. The quantity X_j^{M} (resp. X_j^{P}) equals 0 if the individual carries allele 0 on its maternally (resp. paternally) inherited copy of locus j, and equals 1 otherwise.

The frequencies of allele 1 at locus j on the maternally and paternally inherited genes (averages of X_j^{M} and X_j^{P} over the whole population) are denoted p_j^{M} and p_j^{P}. Finally, $p_j = \left(p_j^{\mathrm{M}} + p_j^{\mathrm{P}}\right)/2$ is the frequency of allele 1 at locus j in the whole population.

Genetic associations. Throughout the paper, index i will denote a locus affecting the selfing rate of individuals, while indices j and k will denote loci affecting fitness. Following Barton and Turelli (1991) and Kirkpatrick et al. (2002), we define the centered variables:

$$\zeta_i^{\mathrm{M}} = \sigma_i^{\mathrm{M}} - \overline{\sigma_i^{\mathrm{M}}}, \quad \zeta_i^{\mathrm{P}} = \sigma_i^{\mathrm{P}} - \overline{\sigma_i^{\mathrm{P}}}, \tag{2}$$

$$\zeta_j^{\mathrm{M}} = X_j^{\mathrm{M}} - p_j^{\mathrm{M}}, \quad \zeta_j^{\mathrm{P}} = X_j^{\mathrm{P}} - p_j^{\mathrm{P}}, \tag{3}$$

where $\overline{\sigma_i^{\mathrm{M}}}$ and $\overline{\sigma_i^{\mathrm{P}}}$ are the averages of σ_i^{M} and σ_i^{P} over the whole population. The genetic association between the sets $\mathbb{U}$ and $\mathbb{V}$ of loci present in the maternally and paternally derived genome of an individual is defined as:

$$D_{\mathbb{U},\mathbb{V}} = \mathrm{E}\left[\zeta_{\mathbb{U},\mathbb{V}}\right] \tag{4}$$

where E stands for the average over all individuals in the population, and with:

$$\zeta_{\mathbb{U},\mathbb{V}} = \left(\prod_{x\in\mathbb{U}} \zeta_x^{\mathrm{M}}\right)\left(\prod_{y\in\mathbb{V}} \zeta_y^{\mathrm{P}}\right). \tag{5}$$

For example, $D_{j,j} = \mathrm{E}\left[\left(X_j^{\mathrm{M}} - p_j^{\mathrm{M}}\right)\left(X_j^{\mathrm{P}} - p_j^{\mathrm{P}}\right)\right]$ is a measure of the departure from Hardy-Weinberg equilibrium at locus j, while $D_{\emptyset,jk} = \mathrm{E}\left[\left(X_j^{\mathrm{P}} - p_j^{\mathrm{P}}\right)\left(X_k^{\mathrm{P}} - p_k^{\mathrm{P}}\right)\right]$ measures the linkage disequilibrium between loci j and k on paternally derived haplotypes. Finally, $\tilde{D}_{\mathbb{U},\mathbb{V}}$ is defined as $\left(D_{\mathbb{U},\mathbb{V}} + D_{\mathbb{V},\mathbb{U}}\right)/2$, and $\tilde{D}_{\mathbb{U},\emptyset}$ will be denoted $\tilde{D}_{\mathbb{U}}$.

Using these notations, the variance in selfing rate in the population can be

written as:

$$V_\sigma = \mathrm{E}\left[\left(\sum_i \left(\zeta_i^{\mathrm{M}} + \zeta_i^{\mathrm{P}}\right)\right)^2\right]. \tag{6}$$

Ignoring genetic associations between different loci affecting the selfing rate, this becomes:

$$V_\sigma \approx 2\sum_i \left(\tilde{D}_{ii} + D_{i,i}\right). \tag{7}$$

General expression for fitness, and special cases. The fitness of an individual divided by the population mean fitness $\overline{W}$ can be expressed in terms of "selection coefficients" $a_{\mathbb{U},\mathbb{V}}$ representing the effect of selection acting on the sets $\mathbb{U}$ and $\mathbb{V}$ of loci (Barton and Turelli, 1991; Kirkpatrick et al., 2002):

$$\frac{W}{\overline{W}} = 1 + \sum_{\mathbb{U},\mathbb{V}} a_{\mathbb{U},\mathbb{V}} \left(\zeta_{\mathbb{U},\mathbb{V}} - D_{\mathbb{U},\mathbb{V}}\right). \tag{8}$$

Throughout the paper, we assume no effect of the sex-of-origin of genes on fitness, so that $a_{\mathbb{U},\mathbb{V}} = a_{\mathbb{V},\mathbb{U}}$. The coefficient $a_{j,\emptyset} = a_{\emptyset,j}$ will be denoted a_j and represents selection for allele 1 at locus j. The coefficient $a_{j,j}$ represents the effect of dominance at locus j, while $a_{jk,\emptyset}$ and $a_{j,k}$ represent cis and trans epistasis between loci j and k. Coefficients $a_{jk,j}$ and $a_{jk,jk}$ respectively correspond to additive-by-dominance and dominance-by-dominance epistatic interactions between loci j and k, measured as deviations from additivity. Throughout the paper, we will assume that selection is weak, all $a_{\mathbb{U},\mathbb{V}}$ being of order ϵ (where ϵ is a small term), and derive general expressions for inbreeding depression and the strength of selection for selfing to leading order in $a_{\mathbb{U},\mathbb{V}}$ coefficients. Results for any particular fitness function can then be obtained by computing the corresponding expressions for $a_{\mathbb{U},\mathbb{V}}$ coefficients. We will consider three examples of fitness function that have been used in previous papers, and lead to different properties of the

three forms of epistasis mentioned above. Approximate expressions for $a_{\mathbb{U},\mathbb{V}}$ coefficients under these fitness functions are computed in Supplementary File S1.

Uniformly deleterious alleles. Our first example corresponds to the case where allele 1 at each fitness locus j is deleterious, with selection and dominance coefficients s and h. Epistatic interactions occur between pairs of loci, and are decomposed into additive-by-additive (e_{axa}), additive-by-dominance (e_{axd}) and dominance-by-dominance (e_{dxd}) epistasis (see Supplementary Figure S1 for an interpretation of these terms). We assume multiplicative effects of epistatic components on fitness W (*i.e.*, additive effects on $\log W$), so that:

$$W = (1 - h\,s)^{n_{\mathrm{he}}} (1 - s)^{n_{\mathrm{ho}}} (1 + e_{\mathrm{axa}})^{n_2} (1 + e_{\mathrm{axd}})^{n_3} (1 + e_{\mathrm{dxd}})^{n_4} \tag{9}$$

where n_{he} and n_{ho} are the numbers of loci at which a deleterious allele is present in the heterozygous (n_{he}) or homozygous (n_{ho}) state, while n_2, n_3 and n_4 are the numbers of interactions between 2, 3 and 4 deleterious alleles at two different loci, given by:

$$n_2 = \frac{1}{2} n_{\mathrm{he}} (n_{\mathrm{he}} - 1) + 2 n_{\mathrm{he}} n_{\mathrm{ho}} + 2 n_{\mathrm{ho}} (n_{\mathrm{ho}} - 1)\,, \tag{10}$$

$$n_3 = n_{\mathrm{he}} n_{\mathrm{ho}} + 2 n_{\mathrm{ho}} (n_{\mathrm{ho}} - 1)\,, \tag{11}$$

$$n_4 = \frac{1}{2} n_{\mathrm{ho}} (n_{\mathrm{ho}} - 1)\,. \tag{12}$$

For example, n_2 is given by the number of pairs of heterozygous loci in the genome ($n_{\mathrm{he}} (n_{\mathrm{he}} - 1)/2$), plus twice the number of pairs involving one heterozygous locus and one homozygous locus for the deleterious allele ($n_{\mathrm{he}} n_{\mathrm{ho}}$), plus four times the number of pairs of homozygous loci for the deleterious allele ($n_{\mathrm{ho}} (n_{\mathrm{ho}} - 1)/2$). In such models with fixed epistasis and possibly large numbers of loci, combinations of mutations

quickly become advantageous when epistasis is positive, in which case they sweep through the population. We therefore focused on cases where e_{axa}, e_{axd} and e_{dxd} are negative, and will assume throughout that deleterious alleles stay at low frequencies in the population (p_j remains small). As shown in Supplementary File S1, equation 9 leads to $a_{jk} = a_{j,k} \approx e_{\mathrm{axa}}$, $a_{jk,j} \approx e_{\mathrm{axd}}$ and $a_{jk,jk} \approx e_{\mathrm{dxd}}$, while the strength of directional selection at each locus (a_j) is affected by e_{axa} and the effective dominance ($a_{j,j}$) is affected by e_{axd}. Because epistatic coefficients are the same for all pairs of loci, equation 9 leads to a situation where the variances of a_{jk}, $a_{jk,j}$ and $a_{jk,jk}$ over pairs of loci equal zero, while their mean values may depart from zero.

Charlesworth et al. (1991) explored the effect of synergistic epistasis (measured by a parameter β) on inbreeding depression, using a fitness function that imposes relations between h, e_{axa}, e_{axd} and e_{dxd}. As explained in Supplementary File S1, their fitness function (equation 2 in Charlesworth et al., 1991) is equivalent to setting $e_{\mathrm{axa}} = -\beta h^2$, $e_{\mathrm{axd}} = -\beta h (1-2h)$ and $e_{\mathrm{dxd}} = -\beta (1-2h)^2$ in our equation 9.

Gaussian stabilizing selection. Our second fitness function corresponds to stabilizing selection acting on an arbitrary number n of quantitative traits, with a symmetrical, Gaussian-shaped fitness function. The general model is the same as in Abu Awad and Roze (2018): $r_{\alpha j}$ denotes the effect of allele 1 at locus j on trait α, and we assume that the different loci have additive effects on traits:

$$g_\alpha = \sum_j r_{\alpha j} \left(X_j^{\mathrm{M}} + X_j^{\mathrm{P}} \right) \tag{13}$$

where g_α is the value of trait α in a given individual (note that $g_\alpha = 0$ for all traits in an individual carrying allele 0 at all loci). We assume that the values of $r_{\alpha j}$ for all

loci and traits are sampled from the same distribution with mean zero and variance a^2. The fitness of individuals is given by:

$$W = \exp\left[-\frac{\sum_{\alpha=1}^{n} {g_\alpha}^2}{2V_\mathrm{s}}\right] \tag{14}$$

where V_s represents the strength of selection. According to equation 14, the optimal value of each trait is zero. We assume that $r_{\alpha j}/\sqrt{V_\mathrm{s}}$ is small, so that selection is weak at each locus. This model generates distributions of fitness effects of mutations and of pairwise epistatic effects on fitness (the average value of epistasis being zero), while deleterious alleles have a dominance coefficient close to 1/4 in an optimal genotype (Martin and Lenormand, 2006b; Martin et al., 2007; Manna et al., 2011). In a population at equilibrium, equations 13 and 14 lead to $a_j \approx -\sum_{\alpha=1}^{n} {r_{\alpha j}}^2 (1-2p_j)/(2V_\mathrm{s})$ (*i.e.*, the rarer allele at locus j is disfavored), $a_{j,j} \approx -\sum_{\alpha=1}^{n} {r_{\alpha j}}^2/V_\mathrm{s}$ and $a_{jk} = a_{j,k} \approx -\sum_{\alpha=1}^{n} r_{\alpha j} r_{\alpha k}/V_\mathrm{s}$, while $a_{jk,j}$ and $a_{jk,jk}$ are smaller in magnitude (see Supplementary File S1). This scenario thus generates a situation where additive-by-additive epistasis ($a_{jk} = a_{j,k}$) is zero on average (because the average of $r_{\alpha j}$ is zero) but has a positive variance among pairs of loci, while additive-by-dominance and dominance-by-dominance epistasis are negligible. As in the previous example, we will generally assume that the deleterious allele at each locus j (allele 1 if $p_j < 0.5$, allele 0 if $p_j > 0$) stays rare in the population, by assuming that $(1-2p_j)^2$ is close to 1; this is also true in the next example.

Non-Gaussian stabilizing selection. The last example we examined is a generalization of the fitness function given by equation 14, in order to introduce a coefficient Q affecting the shape of the fitness peak (e.g., Martin and Lenormand, 2006a; Tenaillon

et al., 2007; Gros et al., 2009; Roze and Blanckaert, 2014; Abu Awad and Roze, 2018):

$$W = \exp\left[-\left(\frac{d}{\sqrt{2V_{\mathrm{s}}}}\right)^{Q}\right], \tag{15}$$

where $d = \sqrt{\sum_{\alpha=1}^{n} {g_\alpha}^2}$ is the Euclidean distance from the optimum in phenotypic space. The fitness function is thus Gaussian when $Q = 2$, while $Q > 2$ leads to a flatter fitness peak around the optimum. The expressions for $a_{\mathbb{U},\mathbb{V}}$ coefficients derived in Supplementary File S1 show that the variances of $a_{jk} = a_{j,k}$, $a_{jk,j}$ and $a_{jk,jk}$ over pairs of loci have the same order of magnitude, and that additive-by-additive epistasis ($a_{jk} = a_{j,k}$) is zero on average, while additive-by-dominance and dominance-by-dominance epistasis ($a_{jk,j}$, $a_{jk,jk}$) are negative on average when $Q > 2$. Note that $Q > 2$ also generates higher-order epistatic interactions (involving more than two loci); however, we did not compute expressions for these terms.

Quasi-linkage equilibrium (QLE) approximation. Using the general expression for fitness given by equation 8, the change in the mean selfing rate per generation can be expressed in terms of genetic associations between loci affecting the selfing rate and loci affecting fitness. Expressions for these associations can then be computed using general methods to derive recursions on allele frequencies and genetic associations (Barton and Turelli, 1991; Kirkpatrick et al., 2002). For this, we decompose the life cycle into two steps: selection corresponds to the differential contribution of individuals due to differences in overall fecundity and/or survival rates (W), while reproduction corresponds to gamete production and fertilization (involving either selfing or outcrossing). Associations measured after selection (that is, weighting each parent by its relative fitness) will be denoted $D'_{\mathbb{U},\mathbb{V}}$, while associations after reproduction (among

offspring) will be denoted $D''_{\mathbb{U},\mathbb{V}}$. Assuming that "effective recombination rates" (that is, recombination rates multiplied by outcrossing rates) are sufficiently large relative to the strength of selection, genetic associations equilibrate rapidly relative to the change in allele frequencies due to selection. In that case, associations can be expressed in terms of allele frequencies by computing their values at equilibrium, for given allele frequencies (e.g., Barton and Turelli, 1991; Nagylaki, 1993). Note that when allele frequencies at fitness loci have reached an equilibrium (for example, at mutation-selection balance), one does not need to assume that the selection coefficients $a_{\mathbb{U},\mathbb{V}}$ are small relative to effective recombination rates for the QLE approximation to hold, but only that changes in allele frequencies due to the variation in the selfing rate between individuals are small. We will thus assume that the variance in the selfing rate in the population V_σ stays small (and therefore, the genetic variance contributed by each locus affecting the selfing rate is also small), and compute expressions to the first order in V_σ. This is equivalent to the assumption that alleles at modifier loci have small effects, commonly done in modifier models.

Individual-based simulations. In order to verify our analytical results, individual-based simulations were run using two C++ programs, one with uniformly deleterious alleles with fixed epistatic effects (equation 9) and the other with stabilizing selection on n quantitative traits (equation 14). Both are described in Supplementary File S5 (and are available from Dryad). Both programs represent a population of N diploid individuals with discrete generations, the genome of each individual consisting of two copies of a linear chromosome with map length R Morgans. In the first program (fixed epistasis), deleterious alleles occur at rate U par haploid genome per generation at

an infinite number of possible sites along the chromosome. A locus with an infinite number of possible alleles, located at the mid-point of the chromosome controls the selfing rate of the individual. In the program representing stabilizing selection, each chromosome carries ℓ equidistant biallelic loci affecting the n traits under selection (as in Abu Awad and Roze, 2018). The selfing rate is controlled by $\ell_\sigma = 10$ additive loci evenly spaced over the chromosome, each with an infinite number of possible alleles (the selfing rate being set to zero if the sum of allelic values at these loci is negative, and one if the sum is larger than one). In both programs, mutations affecting the selfing rate occur at rate $U_{\text{self}} = 10^{-3}$ per generation, the value of each mutant allele at a selfing modifier locus being drawn from a Gaussian distribution with standard deviation σ_{self} centered on the allele value before mutation. The selfing rate is set to zero during an initial burn-in period (set to 20,000 generations) after which mutations are introduced at selfing modifier loci.

RESULTS

Effects of epistasis on inbreeding depression. We first explore the effects of epistasis on inbreeding depression, assuming that the selfing rate is fixed. Throughout the paper, inbreeding depression δ is classically defined as:

$$\delta = 1 - \frac{\overline{W}^{\text{self}}}{\overline{W}^{\text{out}}} \tag{16}$$

where $\overline{W}^{\text{self}}$ and $\overline{W}^{\text{out}}$ are the mean fitnesses of offspring produced by selfing and by outcrossing, respectively (e.g., Lande and Schemske, 1985). In Supplementary File S2, we show that a general expression for δ in terms of one- and two-locus selection

coefficients, in a randomly mating population ($\sigma = 0$) is given by:

$$\delta \approx -\frac{1}{2}\sum_{j} a_{j,j}\, p_j q_j - \frac{1}{2}\sum_{j<k} a_{jk,jk}\left[1 - 2\rho_{jk}\left(1-\rho_{jk}\right)\right] p_j q_j p_k q_k - \sum_{j<k} c_{jk}\, \tilde{D}_{jk} \tag{17}$$

where the sums are over all loci affecting fitness, and with:

$$c_{jk} = a_{j,k} + \left[a_{jk,j}\left(1-2p_j\right) + a_{jk,k}\left(1-2p_k\right)\right]\left(1-\rho_{jk}\right), \tag{18}$$

ρ_{jk} being the recombination rate between loci j and k. With arbitrary selfing, and assuming all $\rho_{jk} \approx 1/2$, equation 17 generalizes to:

$$\delta \approx -\frac{1}{2}\sum_{j} a_{j,j}\left(1+F\right) p_j q_j - \frac{1}{4}\sum_{j<k} a_{jk,jk}\left[\left(1+F\right)^2 + G_{jk}\right] p_j q_j p_k q_k \tag{19}$$

with several higher-order terms depending on genetic associations between loci generated by epistatic interactions ($\tilde{D}_{jk}$, $\tilde{D}_{j,k}$, $\tilde{D}_{jk,j}$, see equation B17 in Supplementary File S2 for the complete expression). The term F in equation 19 corresponds to the inbreeding coefficient (probability of identity by descent between the maternal and paternal copy of a gene), given by:

$$F = \frac{\sigma}{2-\sigma} \tag{20}$$

at equilibrium, while G_{jk} is the identity disequilibrium between loci j and k (Weir and Cockerham, 1973), given by:

$$G_{jk} = \phi_{jk} - F^2, \quad \text{with} \quad \phi_{jk} = \frac{\sigma}{2-\sigma}\frac{2-\sigma-2\left(2-3\sigma\right)\rho_{jk}\left(1-\rho_{jk}\right)}{2-\sigma\left[1-2\rho_{jk}\left(1-\rho_{jk}\right)\right]} \tag{21}$$

(ϕ_{jk} is the joint probability of identity by descent at loci j and k). Under free recombination ($\rho_{jk} = 1/2$), it simplifies to:

$$G_{jk} = \frac{4\sigma\left(1-\sigma\right)}{\left(4-\sigma\right)\left(2-\sigma\right)^2}, \tag{22}$$

which will be denoted G hereafter. Given that G_{jk} is only weakly dependent on ρ_{jk}, G_{jk} should be close to G for most pairs of loci when the genome map length is not too small.

Uniformly deleterious alleles. When fitness is given by equation 9, from equation 19 and using the expressions for the $a_{\mathbb{U},\mathbb{V}}$ coefficients given in Supplementary File S1 we find:

$$\delta \approx \frac{1}{2}\left[s\left(1-2h\right)-2e_{\mathrm{axd}}\,n_{\mathrm{d}}\right]\left(1+F\right)n_{\mathrm{d}}-\frac{e_{\mathrm{dxd}}}{8}\left[\left(1+F\right)^{2}+G\right]n_{\mathrm{d}}^{2} \tag{23}$$

where $n_{\mathrm{d}} = \sum_j p_j$ is the average number of deleterious alleles per haploid genome. Equation 24 assumes that deleterious alleles stay rare in the population (so that terms in ${p_j}^2$ may be neglected). As explained in Supplementary File S2, the expression:

$$\delta \approx 1-\exp\left[-\frac{1}{2}\left[s\left(1-2h\right)-2e_{\mathrm{axd}}\,n_{\mathrm{d}}\right]\left(1+F\right)n_{\mathrm{d}}+\frac{e_{\mathrm{dxd}}}{8}\left[\left(1+F\right)^{2}+G\right]n_{\mathrm{d}}^{2}\right] \tag{24}$$

(obtained by assuming that the effects on inbreeding depression of individual loci and their interactions do multiply, rather than sum) provides more accurate results for parameter values leading to high inbreeding depression. Equation 24 yields the classical expression $\delta \approx 1-\exp\left[-U\left(1-2h\right)/\left(2h\right)\right]$ in the absence of epistasis and under random mating (e.g., Charlesworth and Charlesworth, 1987).

Equations 23 and 24 only depend on the mean number of deleterious alleles n_{d} (and not on recombination rates between selected loci) because the effects of genetic associations between loci on δ have been neglected (as they generate higher-order terms, whose effects should remain small in most cases), and because G_{jk} was approximated by G. The equilibrium value of n_{d} can be obtained by solving

$$\Delta_{\mathrm{sel}}\,n_{\mathrm{d}}+U=0 \tag{25}$$

where $\Delta_{\text{sel}} n_{\text{d}} = \sum_j \Delta_{\text{sel}} p_j$ is the change in n_{d} due to selection and U is the deleterious mutation rate per haploid genome. From equation B26 in Supplementary File S2, we have to the first order in the selection coefficients (and assuming that deleterious alleles stay rare):

$$\begin{aligned}\Delta_{\text{sel}} p_j \approx a_j \left(1+F\right) p_j + a_{j,j}\, F\, p_j + \sum_{k\neq j} a_{jk,k} \left[F\left(1+F\right) + G_{jk}\right] p_j p_k \\ + \sum_{k\neq j} a_{jk,jk} \left[F^2 + G_{jk}\right] p_j p_k\end{aligned} \tag{26}$$

simplifying to $a_j p_j$ under random mating. The first term of equation 26 represents the effect of directional selection against deleterious alleles ($a_j < 0$), which is increased by selfing due to the higher variance between individuals generated by homozygosity (by a factor $1 + F$). The other terms represent additional effects of dominance and epistatic terms involving dominance arising when $\sigma > 0$ (that is, when the frequency of homozygous mutants is not negligible). Summing over loci and using the expressions for $a_{\mathbb{U},\mathbb{V}}$ coefficients given in Supplementary File S1, one obtains:

$$\begin{aligned}\Delta_{\text{sel}} n_{\text{d}} \approx -s \left[h + \left(1-h\right) F\right] n_{\text{d}} + 2 e_{\text{axa}} \left(1+F\right) n_{\text{d}}^2 \\ + e_{\text{axd}} \left[F\left(3+F\right) + G\right] n_{\text{d}}^2 + e_{\text{dxd}} \left(F^2 + G\right) n_{\text{d}}^2\end{aligned} \tag{27}$$

that can be used with equation 25 to obtain the equilibrium value of n_{d} (note that the term in e_{axa} in equation 27 stems from the term in a_j in equation 26, while part of the term in e_{axd} stems from the term in $a_{j,j}$).

Equation 27 shows that, for non-random mating, negative values of e_{axa}, e_{axd} or e_{dxd} reduce the mean number of deleterious alleles at equilibrium, thereby reducing inbreeding depression (the effects of e_{axd} and e_{dxd} on the equilibrium value of n_{d} disappear when mating is random, as $F = G = 0$ in this case). As shown by equation 24, negative values of e_{axd} and e_{dxd} also directly increase inbreeding depression (even

under random mating), by decreasing the fitness of homozygous offspring. Figures 1A–C compare the predictions obtained from equations 24 and 27 with simulation results, testing the effect of each epistatic component separately. Negative e_{axa} reduces inbreeding depression by lowering the frequency of deleterious alleles in the population (equation 27, Figure 1A); furthermore, it reduces the purging effect of selfing, so that inbreeding depression may remain constant or even slightly increase as the selfing rate increases. When the selfing rate is low, e_{axd} and e_{dxd} have little effect on the mean number of deleterious alleles n_{d}, and the main effect of negative e_{axd} and e_{dxd} is to increase inbreeding depression by decreasing the fitness of homozygous offspring (equation 24, Figures 1B–C). As selfing increases, this effect becomes compensated by the enhanced purging caused by negative e_{axd} and e_{dxd} (equation 27). Figure 1D shows the results obtained using Charlesworth et al.'s (1991) fitness function, yielding $e_{\mathrm{axa}} = -\beta h^2$, $e_{\mathrm{axd}} = -\beta h (1-2h)$ and $e_{\mathrm{dxd}} = -\beta (1-2h)^2$. Remarkably, the increased purging caused by negative epistasis almost exactly compensates the decreased fitness of homozygous offspring, so that inbreeding depression is only weakly affected by epistasis in this particular model; this result is also observed for different values of s, U and h (Supplementary Figures S2 – S4). The discrepancies between analytical and simulation results in Figure 1 likely stem from the effects of genetic associations, which are neglected in equations 24 and 27 (e.g., Roze, 2015). As shown by Supplementary Figures S2 – S4, these discrepancies become more important as the strength of epistasis increases (relative to s), as the mutation rate U increases and as dominance h decreases.

Gaussian stabilizing selection. An expression for inbreeding depression under Gaussian stabilizing selection (equation 14) is given in Abu Awad and Roze (2018). As shown in Supplementary File S2, this expression can be recovered from our general expression for δ in terms of $a_{\mathbb{U},\mathbb{V}}$ coefficients. Because the average epistasis is zero under Gaussian selection (e.g., Martin et al., 2007), inbreeding depression is only affected by the variance in epistasis, whose main effect is to generate linkage disequilibria that increase the frequency of deleterious alleles (see also Phillips et al., 2000) and thus increase δ. As shown by Abu Awad and Roze (2018), a different regime is entered above a threshold selfing rate when the mutation rate U is sufficiently large, in which epistatic interactions tend to lower inbreeding depression (see also Lande and Porcher, 2015).

Non-Gaussian stabilizing selection. Expressions for $a_{\mathbb{U},\mathbb{V}}$ coefficients under the more general fitness function given by equation 15 (Supplementary File S1) show that a "flatter-than-Gaussian" fitness peak ($Q > 2$) generates negative dominance-by-dominance epistasis ($a_{jk,jk} < 0$), increasing inbreeding depression (by contrast, the first term of equation 17 representing the effect of dominance is not affected by Q, as the effects of Q on $a_{j,j}$ and on $p_j q_j$ cancel out). In the absence of selfing, and neglecting the effects of genetic associations among loci, one obtains (see Supplementary File S2 for derivation):

$$\delta \approx 1 - \exp\left[-U\left(1 + \frac{Q-2}{8}\right)\right] \tag{28}$$

where the term in $(Q-2)/8$ is generated by the term in $a_{jk,jk}$ in equation 17. Although this expression differs from equation 29 in Abu Awad and Roze (2018) — that was obtained using a different method — both results are quantitatively very similar as long as Q is not too large (roughly, $Q < 6$). Generalizations of equation 28 to arbitrary

σ, and including the effects of pairwise associations between loci (for $\sigma = 0$) are given in Supplementary File S2 (equations B40 and B54).

Evolution of selfing in the absence of epistasis. Before exploring the effects of epistasis on selection for selfing, we first derive a general expression for the strength of indirect selection for selfing in the absence of epistasis (that is, ignoring the effects of coefficients a_{jk}, $a_{j,k}$, $a_{jk,j}$ and $a_{jk,jk}$ of the fitness function). In Supplementary File S3, we show that the change in the mean selfing rate $\overline{\sigma}$ per generation can be decomposed into three terms:

$$\Delta\overline{\sigma} = \Delta_{\text{auto}}\overline{\sigma} + \Delta_{\text{depr}}\overline{\sigma} + \Delta_{\text{purge}}\overline{\sigma} \tag{29}$$

with:

$$\Delta_{\text{auto}}\overline{\sigma} \approx \frac{1-\kappa}{1-\kappa\,\overline{\sigma}}\frac{V'_\sigma}{2}, \tag{30}$$

$$\Delta_{\text{depr}}\overline{\sigma} = 2\sum_{i,j} a_{j,j}\,\tilde{D}_{ij,j}, \tag{31}$$

$$\Delta_{\text{purge}}\overline{\sigma} = 2\sum_{i,j} a_j\left(\tilde{D}_{ij} + \tilde{D}_{i,j}\right) \tag{32}$$

where the sums are over all loci i affecting the selfing rate and all loci j affecting fitness. The term $\Delta_{\text{auto}}\overline{\sigma}$ represents selection for increased selfing rates due to the automatic transmission advantage associated with selfing (Fisher, 1941). It is proportional to the variance in the selfing rate after selection V'_σ, and vanishes when pollen discounting is complete ($\kappa = 1$). The second term corresponds to the effect of inbreeding depression. It depends on coefficients $a_{j,j}$, representing the effect of dominance at loci affecting fitness; in particular, $a_{j,j} < 0$ when the average fitness of the two homozygotes at locus j is lower than the fitness of heterozygotes (which is the case when the deleterious allele at locus j is recessive or partially recessive). It also depends on associations

$\tilde{D}_{ij,j}$ that are shown to be positive at QLE, reflecting the fact that alleles increasing the selfing rate tend to be present on more homozygous backgrounds. Finally, the last term depends on coefficients a_j representing directional selection for allele 1 at locus j, and associations $\tilde{D}_{ij}$ and $\tilde{D}_{i,j}$ which are positive when alleles increasing the selfing rate at locus i tend to be associated with allele 1 at locus j, either on the same or on the other haplotype. This term is generally positive (favoring increased selfing rates), representing the fact that alleles coding for higher selfing increase the efficiency of selection at selected loci (by increasing homozygosity), and thus tend to be found on better purged genetic backgrounds, as explained in the Introduction (we show in Supplementary File S3 that $\tilde{D}_{ij}$ and $\tilde{D}_{i,j}$ are also generated by other effects involving the identity disequilibrium between loci i and j, when $0 < \overline{\sigma} < 1$).

The variance in the selfing rate after selection V'_σ, and the associations $\tilde{D}_{ij,j}$, $\tilde{D}_{ij}$ and $\tilde{D}_{i,j}$ can be expressed in terms of V_σ and of allele frequencies using the QLE approximation described in the Methods. The derivations and expressions obtained for arbitrary values of $\overline{\sigma}$ can be found in Supplementary File S3 (equations C31, C47, C48, C55 and C64), and generalize the results given by Epinat and Lenormand (2009) in the case of strong discounting ($\kappa \approx 1$). When the mean selfing rate in the population approaches zero, one obtains:

$$V'_\sigma \approx V_\sigma, \quad \tilde{D}_{ij,j} \approx \frac{1}{2}\tilde{D}_{ii}\, p_j q_j, \tag{33}$$

$$\tilde{D}_{ij} \approx \frac{1}{2}\frac{a_j + a_{j,j}\,(1-2p_j)}{\rho_{ij} - a_j\,(1-2p_j)\,(1-\rho_{ij})}\tilde{D}_{ii}\, p_j q_j, \quad \tilde{D}_{i,j} \approx 0. \tag{34}$$

Using the fact that $V_\sigma = 2\sum_i \tilde{D}_{ii}$ under random mating (equation 7), equations 30 – 34 yield, for $\overline{\sigma} \approx 0$:

$$\Delta_{\text{auto}}\overline{\sigma} \approx \frac{1-\kappa}{2}\, V_\sigma, \quad \Delta_{\text{depr}}\overline{\sigma} \approx -\delta\, V_\sigma, \tag{35}$$

where $\delta = -\left(\sum_j a_{j,j}\, p_j q_j\right)/2$ is inbreeding depression, neglecting the effect of interactions between selected loci (see equation 17), while

$$\Delta_{\text{purge}}\overline{\sigma} \approx \sum_j \left[\mathcal{E}\left[\frac{1}{\rho_{ij} - a_j\,(1-2p_j)\,(1-\rho_{ij})}\right] a_j\,[a_j + a_{j,j}\,(1-2p_j)]\,p_j q_j\right] \frac{V_\sigma}{2} \quad (36)$$

where the sum is over all loci j affecting fitness, and where $\mathcal{E}$ is the average over all loci i affecting the selfing rate. Because $\Delta_{\text{purge}}\overline{\sigma}$ is of second order in the selection coefficients (a_j, $a_{j,j}$), it will generally be negligible relative to $\Delta_{\text{depr}}\overline{\sigma}$ (which is of first order in $a_{j,j}$), in which case selfing can increase if $\delta < (1-\kappa)/2$ (Charlesworth, 1980). When $\overline{\sigma} > 0$, $\Delta_{\text{depr}}\overline{\sigma}$ is not simply given by $-\delta\, V_\sigma$ (in particular, it also depends on the rate of pollen discounting and on identity disequilibria between loci affecting the selfing rate and loci affecting fitness, as shown by equation C31 in Supplementary File S3), but it is possible to show that $\Delta_{\text{depr}}\overline{\sigma}$ tends to decrease in magnitude as $\overline{\sigma}$ increases (while $\Delta_{\text{auto}}\overline{\sigma}$ becomes stronger as $\overline{\sigma}$ increases), leading to the prediction that $\overline{\sigma} = 0$ and $\overline{\sigma} = 1$ should be the only evolutionarily stable selfing rates (Lande and Schemske, 1985). As shown by equation 36, the relative importance of $\Delta_{\text{purge}}\overline{\sigma}$ should increase when the strength of directional selection (a_j) increases, when deviations from additivity ($a_{j,j}$) are weaker and when linkage among loci is tighter.

In Supplementary File S3, we show that equation 36 can be expressed in terms of the increase in mean fitness caused by a single generation of selfing. In particular, if we imagine an experiment where a large pool of selfed offspring and a large pool of outcrossed offspring are produced from the same pool of parents (sampled from a randomly mating population), and if these offspring are allowed to reproduce (in proportion to their fitness, and by random mating within each pool), one can show that the mean fitness of the offspring of selfed individuals will be increased relative

to the offspring of outcrossed individuals (due to purging), by an amount approximately equal to $P = \sum_j a_j \left[a_j + a_{j,j}\left(1 - 2p_j\right)\right] p_j q_j$. Therefore, when linkage between loci affecting selfing and selected loci is not too tight (so that the term in a_j in the denominator of equation 36 may be neglected), $\Delta_{\text{purge}}\overline{\sigma}$ is approximately $P\,V_\sigma/\left(2\rho_{\text{h}}\right)$, where ρ_h is the harmonic mean recombination rate over all pairs of loci i and j, where i affects the selfing rate and j affects fitness (in the case of freely recombining loci, we thus have $\Delta_{\text{purge}}\overline{\sigma} \approx P\,V_\sigma$).

Uniformly deleterious alleles. In the case where allele 1 at each fitness locus is deleterious with selection and dominance coefficients s and h (and assuming that $p_j \ll 1$) we have $a_j \approx -sh$ and $a_{j,j} \approx -s\left(1-2h\right)$, while $p_j q_j \approx u/\left(sh\right)$ at mutation-selection balance (where u is the mutation rate per locus). In that case, equation 36 simplifies to:

$$\Delta_{\text{purge}}\overline{\sigma} \approx \mathcal{E}\left[\frac{1}{\rho_{ij} + sh\left(1-\rho_{ij}\right)}\right] s\left(1-h\right) U \frac{V_\sigma}{2} \tag{37}$$

where U is the deleterious mutation rate per haploid genome and $\mathcal{E}$ is now the average over all pairs of loci i and j (where locus i affects the selfing rate while locus j affects fitness). Figure 2A compares the predictions obtained from equations 35 and 37 with simulation results, in the absence of pollen discounting ($\kappa = 0$), and when alleles affecting the selfing rate have weak effects ($\sigma_{\text{self}} = 0.01$). Simulations confirm that selfing may evolve when inbreeding depression is higher than 0.5 (due to the effect of $\Delta_{\text{purge}}\overline{\sigma}$), provided that the fitness effect of deleterious alleles is sufficiently strong. The prediction for the case of unlinked loci (obtained by setting $\rho_{ij} = 0.5$ in equation 37) actually gives a closer match to the simulation results than the result obtained by integrating equation 37 over the genetic map. This may stem from the

fact that equation 37 overestimates the effect of tightly linked loci (possibly because the QLE approximation becomes inaccurate when ρ_{ij} is of the same order of magnitude as changes in allele frequencies due to selection for selfing). The effect of the size of mutational steps at the modifier locus does not affect the maximum value of inbreeding depression for which selfing can spread, as long as mutations tend to have small effects on the selfing rate (compare Figure 2A and 2B). However, the relative effect of purging (observed for high values of s) becomes more important when selfing can evolve by mutations of large size ($\sigma_{\text{self}} = 0.3$ in Figure 2C, while mutations directly lead to fully selfing individuals in Figure 2D), in agreement with the results obtained by Charlesworth et al. (1990) — note that our approximations break down when selfing evolves by large-effect mutations.

Gaussian and non-Gaussian stabilizing selection. In the case of multivariate Gaussian stabilizing selection acting on n traits coded by biallelic loci with additive effects (equation 14) we have (to the first order in the strength of selection $1/V_{\text{s}}$): $a_j = -\varsigma_j (1 - 2p_j)$ and $a_{j,j} = -2\varsigma_j$, where $\varsigma_j = \sum_{\alpha=1}^{n} {r_{\alpha j}}^2 / (2V_{\text{s}})$ is the fitness effect of a heterozygous mutation at locus j in an optimal genotype. Assuming that polymorphism stays weak at loci coding for the traits under stabilizing selection, so that $(1 - 2p_j)^2 \approx 1$, and using the fact that $p_j q_j \approx u/\varsigma_j$ under random mating (from equation B26, and neglecting interactions between loci), one obtains from equation 36:

$$\Delta_{\text{purge}}\overline{\sigma} \approx \mathcal{E}\left[\frac{3\varsigma_j}{\rho_{ij} + \varsigma_j (1 - \rho_{ij})}\right] U \frac{V_\sigma}{2} \tag{38}$$

which is equivalent to equation 37 when introducing differences in s among loci, with $h = 1/4$ (note that the homozygous effect of mutations at locus j in an optimal

genotype is $\approx 4\varsigma_j$). When neglecting the term in ς_j in the denominator of equation 38, this simplifies to:

$$\Delta_{\text{purge}}\overline{\sigma} \approx \frac{3}{2}\,\frac{\overline{\varsigma}\,U\,V_\sigma}{\rho_{\text{h}}} \tag{39}$$

where $\overline{\varsigma}$ is the average heterozygous effect of mutations on fitness in an optimal genotype, and where ρ_{h} is the harmonic mean recombination rate over all pairs of loci i and j, where i affects the selfing rate and j affects the traits under stabilizing selection. Using the fitness function given by equation 15 (where Q describes the shape of the fitness peak), equation 39 generalizes to:

$$\Delta_{\text{purge}}\overline{\sigma} \approx \frac{3U^2}{\rho_{\text{h}}}\left(\frac{4U}{Q\overline{\varsigma}}\right)^{-\frac{2}{Q}} V_\sigma \tag{40}$$

(see Supplementary File S1), which increases as Q increases in most cases (the derivative of equation 40 with respect to Q is positive as long as $Q < 10.88U/\overline{\varsigma}$). Therefore, for a given value of inbreeding depression and a fixed $\overline{\varsigma}$, a flatter fitness peak tends to increase the relative importance of purging on the spread of selfing mutants in an outcrossing population.

Effects of epistasis on the evolution of selfing. We now extend the previous expressions to include the effect of epistasis between pairs of selected loci. For this, we assume that all selection coefficients $a_{\mathbb{U},\mathbb{V}}$ are of the same order of magnitude (of order ϵ), and derive expressions for the effects of epistatic coefficients a_{jk}, $a_{j,k}$, $a_{jk,j}$ and $a_{jk,jk}$ on the change in mean selfing rate $\overline{\sigma}$ to leading order in ϵ. Because the expressions quickly become cumbersome under partial selfing, we restrict our analysis to the initial spread of selfing in an outcrossing population ($\overline{\sigma} \approx 0$). Figure 3 summarizes the different effects of epistasis, that are detailed below.

As shown in Supplementary File S4, the change in mean selfing rate per generation now writes:

$$\Delta\overline{\sigma} = \Delta_{\text{auto}}\overline{\sigma} + \Delta_{\text{depr}}\overline{\sigma} + \Delta_{\text{LD}}\overline{\sigma} + \Delta_{\text{purge}}\overline{\sigma}\,. \tag{41}$$

As above, $\Delta_{\text{auto}}\overline{\sigma}$ represents the direct transmission advantage of selfing and is still given by equation 35 as $\overline{\sigma}$ tends to zero. The term $\Delta_{\text{depr}}\overline{\sigma}$ corresponds to the effect of inbreeding depression; taking into account epistasis between selected loci, it writes:

$$\begin{aligned}\Delta_{\text{depr}}\overline{\sigma} = {} & 2\sum_{i,j} a_{j,j}\,\tilde{D}_{ij,j} + 2\sum_{i,j<k} a_{jk,jk}\,\tilde{D}_{ijk,jk} \\ & + 2\sum_{i,j<k} a_{j,k}\left(\tilde{D}_{ij,k} + \tilde{D}_{ik,j}\right) + 2\sum_{i,j,k} a_{jk,j}\left(\tilde{D}_{ijk,j} + \tilde{D}_{ij,jk}\right)\end{aligned} \tag{42}$$

As shown in Supplementary File S4, expressing the different associations that appear in equation 42 at QLE, to leading order (and when $\overline{\sigma}$ tends to zero) yields $\Delta_{\text{depr}}\overline{\sigma} = -\delta'\,V_\sigma$, where δ' is inbreeding depression measured after selection, that is, when the parents used to produced selfed and outcrossed offspring contribute in proportion to their fitness (an expression for δ' in terms of allele frequencies and associations between pairs of loci is given by equation B9 in Supplementary File S2). Indeed, what matters for the spread of selfing is the ratio between the mean fitnesses of selfed and outcrossed offspring, taking into account the differential contributions of parents due to their different fitnesses. With epistasis, inbreeding depression is affected by genetic associations between selected loci, and δ' thus depends on the magnitude of those associations after selection. Note that epistasis may also affect inbreeding depression through the effective dominance $a_{j,j}$ and the equilibrium frequency p_j of deleterious alleles (as described earlier), and these effects are often stronger than effects involving genetic associations when epistasis differs from zero on average.

The new term $\Delta_{\text{LD}}\overline{\sigma}$ appearing in equation 41 represents an additional effect of

epistasis (besides its effects on inbreeding depression δ'), and is given by:

$$\Delta_{\mathrm{LD}}\overline{\sigma} = 2\sum_{i,j<k} a_{jk}\,\tilde{D}_{ijk}\,. \tag{43}$$

The association $\tilde{D}_{ijk}$ represents the fact that the linkage disequilibrium D_{jk} between loci j and k (generated by epistasis among those loci) tends to be stronger on haplotypes that also carry an allele increasing the selfing rate at locus i. Indeed, the magnitude of D_{jk} depends on the relative forces of selection generating D_{jk} and recombination breaking it, and selfing affects both processes: by increasing homozygosity, selfing reduces the effect of recombination (e.g., Nordborg, 1997), but it also increases "effective" epistasis, given that when a beneficial combination of alleles is present on one haplotype of an individual, it also tends to be present on the other haplotype due to homozygosity, enhancing the effect of fitness differences between haplotypes. An expression for $\tilde{D}_{ijk}$ at QLE is given in Supplementary File S4, showing that $\tilde{D}_{ijk}$ is generated by all epistatic components (a_{jk}, $a_{j,k}$, $a_{jk,j}$, and $a_{jk,jk}$). However, in the case of stabilizing selection the terms in $a_{jk,j}$, and $a_{jk,jk}$ should vanish when summed over sufficiently large numbers of loci (as explained in Supplementary File S4).

As in the previous section, the term $\Delta_{\mathrm{purge}}\overline{\sigma}$ equals $2\sum_{i,j} a_j\,\tilde{D}_{ij}$ under random mating and represents indirect selection for selfing due to the fact that selfing increases the efficiency of selection against deleterious alleles. At QLE and to the first order in $a_{\mathbb{U},\mathbb{V}}$ coefficients, the linkage disequilibrium $\tilde{D}_{ij}$ is given by (see Supplementary File S4 for derivation):

$$\begin{aligned}\tilde{D}_{ij} \approx \frac{1}{2}\frac{\tilde{D}_{ii}\,p_j q_j}{\rho_{ij} - a_j\,(1-2p_j)\,(1-\rho_{ij})}\Bigg[&a_j + a_{j,j}\,(1-2p_j)\\ &+\sum_k \left[a_{jk,k} + \left[a_{jk,k} + a_{jk,jk}\,(1-2p_j)\right]\left[1-2\rho_{jk}\,(1-\rho_{jk})\right]\right] p_k q_k\Bigg].\end{aligned} \tag{44}$$

The term on the first line of equation 44 is the same as in equation 34, representing

the fact that increased homozygosity at locus j improves the efficiency of selection acting at this locus. Note that epistatic interactions may affect this term (in particular when the average epistasis between selected loci differs from zero) through the selection coefficients a_j and $a_{j,j}$ as well as equilibrium allele frequencies p_j. The term on the second line of equation 44 shows that negative additive-by-dominance ($a_{jk,j}$) or dominance-by-dominance epistasis ($a_{jk,jk}$) between deleterious alleles increase the benefit of selfing: indeed, negative values of $a_{jk,j}$ and $a_{jk,jk}$ increase the magnitude of the negative linkage disequilibrium between alleles increasing selfing and disfavored alleles at loci affecting fitness (allele 1 in the case of uniformly deleterious alleles, or the rarer allele in the case of stabilizing selection). This effect stems from the increased homozygosity of offspring produced by selfing, negative values of $a_{jk,j}$ and $a_{jk,jk}$ increasing the efficiency of selection against deleterious alleles in homozygous individuals.

Uniformly deleterious alleles. Under fixed selection and epistatic coefficients across loci (fitness given by equation 9) and assuming that deleterious alleles stay rare in the population, one obtains for $\Delta_{\mathrm{LD}}\overline{\sigma}$:

$$\Delta_{\mathrm{LD}}\overline{\sigma} \approx \mathcal{E}\left[\frac{e_{\mathrm{axa}}\left(2+{\rho_{jk}}^2\right)+e_{\mathrm{axd}}+\left(e_{\mathrm{axd}}+\frac{1}{2}e_{\mathrm{dxd}}\right)\left[1-2\rho_{jk}\left(1-\rho_{jk}\right)\right]}{\rho_{ijk}-\left(1-\rho_{ijk}\right)\left(a_j+a_k+e_{\mathrm{axa}}\right)}\right]e_{\mathrm{axa}}{n_{\mathrm{d}}}^2\frac{V_\sigma}{2} \tag{45}$$

where $\mathcal{E}$ is the average over all triplets of loci i, j and k, ρ_{ijk} is the probability that at least one recombination event occurs between the three loci i, j and k during meiosis (note that the denominator is approximately ρ_{ijk} when recombination rates are large relative to selection coefficients), and where n_{d} is the mean number of deleterious alleles per haploid genome. Assuming free recombination among all loci ($\rho_{jk} = 1/2$,

$\rho_{ijk} = 3/4$), equation 45 simplifies to:

$$\Delta_{\mathrm{LD}}\overline{\sigma} \approx \frac{e_{\mathrm{axa}}}{6}\left(9e_{\mathrm{axa}} + 6e_{\mathrm{axd}} + e_{\mathrm{dxd}}\right) {n_{\mathrm{d}}}^2 V_{\sigma}, \tag{46}$$

or, using Charlesworth et al.'s (1991) fitness function:

$$\Delta_{\mathrm{LD}}\overline{\sigma} \approx \left[\beta h\left(1+h\right) n_{\mathrm{d}}\right]^2 \frac{V_{\sigma}}{6}. \tag{47}$$

Furthermore, $\Delta_{\mathrm{purge}}\overline{\sigma}$ is given by:

$$\Delta_{\mathrm{purge}}\overline{\sigma} \approx \mathcal{E}\left[\frac{h\left[s\left(1-h\right) - 3e_{\mathrm{axd}}n_{\mathrm{d}} - \left[1-2\rho_{jk}\left(1-\rho_{jk}\right)\right]\left(e_{\mathrm{axd}}+e_{\mathrm{dxd}}\right)n_{\mathrm{d}}\right] - 2e_{\mathrm{axa}}n_{\mathrm{d}}}{\rho_{ij} - \left(1-\rho_{ij}\right)a_j}\right] \times s n_{\mathrm{d}}\frac{V_{\sigma}}{2} \tag{48}$$

simplifying to:

$$\Delta_{\mathrm{purge}}\overline{\sigma} \approx \left[h\left[2s\left(1-h\right) - \left(7e_{\mathrm{axd}} + e_{\mathrm{dxd}}\right)n_{\mathrm{d}}\right] - 4e_{\mathrm{axa}}n_{\mathrm{d}}\right] s n_{\mathrm{d}}\frac{V_{\sigma}}{4} \tag{49}$$

under free recombination.

Figure 4 shows the parameter space (in the κ – δ' plane) in which an initially outcrossing population ($\overline{\sigma} = 0$) evolves towards selfing, in the case of uniformly deleterious alleles. Note that when selfing increased in the simulations (green dots), we always observed that the population evolved towards selfing rates close to 1. Figures 4A–C show that negative e_{axd} or e_{dxd} (the other epistatic components being set to zero) slightly increase the parameter range under which selfing evolves: in particular, selfing can invade for values of inbreeding depression δ' slightly higher than 0.5 in the absence of pollen discounting ($\kappa = 0$). Epistasis has stronger effects when negative e_{axd} and/or e_{dxd} are combined with negative e_{axa}, as shown by Figures 4D–F (we did not test the effect of negative e_{axa} alone, as δ' is greatly reduced in this case unless

e_{axa} is extremely weak). The QLE model (dashed and solid curves) correctly predicts the maximum inbreeding depression δ' for selfing to evolve, as long as this maximum is not too large: high values of δ' indeed imply high values of U, for which the QLE model overestimates the strength of indirect effects (in particular, the model predicts that selfing may evolve under high depression, above the upper parts of the curves in Figures 4D–F, but this was never observed in the simulations). This discrepancy may stem from higher-order associations between selected loci (associations involving 3 or more selected loci), that are neglected in this analysis and may become important when large numbers of mutations are segregating.

In all cases shown in Figure 4, the increased parameter range under which selfing can evolve is predicted to be mostly due to the effect of negative epistasis on $\Delta_{\mathrm{purge}}\overline{\sigma}$, the effect of $\Delta_{\mathrm{LD}}\overline{\sigma}$ remaining negligible. Finally, one can note that the maximum δ' for selfing to evolve is lower with $e_{\mathrm{axa}} = -0.005$, $e_{\mathrm{axd}} = e_{\mathrm{dxd}} = -0.01$ (Figure 4E) than with $e_{\mathrm{axa}} = -0.005$, $e_{\mathrm{axd}} = -0.01$, $e_{\mathrm{dxd}} = 0$ (Figure 4D). This is due to the fact that negative e_{axd} and e_{dxd} have two opposite effects: they increase the effect of selection against homozygous mutations (which increases $\Delta_{\mathrm{purge}}\overline{\sigma}$), but they also increase the strength of inbreeding depression for a given mutation rate U (see Figure 1), decreasing the mean number of deleterious alleles per haplotype n_{d} associated with a given value of δ' (which decreases $\Delta_{\mathrm{purge}}\overline{\sigma}$).

Supplementary Figure S5 shows the effect of the size of mutational steps at the selfing modifier locus, in the absence of epistasis (corresponding to Figure 4A), and with all three components of epistasis being negative (corresponding to Figure 4E). Increasing the size of mutational steps has more effect in the presence of negative epistasis, since negative epistasis increases the purging advantage of alleles coding for

more selfing ($\Delta_{\text{purge}}\overline{\sigma}$), whose effect becomes stronger relative to $\Delta_{\text{auto}}\overline{\sigma}$ and $\Delta_{\text{depr}}\overline{\sigma}$ when modifier alleles have larger effects (as previously shown in Figure 2).

Gaussian and non-Gaussian stabilizing selection. Under stabilizing selection acting on quantitative traits (and assuming that recombination rates are not too small), one obtains:

$$\Delta_{\text{LD}}\overline{\sigma} \approx \mathcal{E}\left[\frac{2+{\rho_{jk}}^2}{\rho_{ijk}}\right]\frac{2U^2}{n}\,V_\sigma, \tag{50}$$

(where n is the number of selected traits) independently of the shape of the fitness peak Q, simplifying to $(6U^2/n)\,V_\sigma$ under free recombination (see Supplementary File S4). Independence from Q stems from the fact that $\Delta_{\text{LD}}\overline{\sigma}$ is proportional to $\sum_{j,k}{a_{jk}}^2 p_j q_j p_k q_k$, while Q has opposite effects on ${a_{jk}}^2$ and on $p_j q_j p_k q_k$ (${a_{jk}}^2$ decreases, while $p_j q_j p_k q_k$ increases as Q increases), which compensate each other exactly in this sum.

Under Gaussian stabilizing selection ($Q = 2$), the coefficients $a_{jk,j}$ and $a_{jk,jk}$ are small relative to the other selection coefficients (as shown in Supplementary File S1), and their effect on $\Delta_{\text{purge}}\overline{\sigma}$ may thus be neglected (in which case $\Delta_{\text{purge}}\overline{\sigma}$ is still given by equation 39). With a flatter fitness peak ($Q > 2$), using the expressions for $a_{jk,j}$ and $a_{jk,jk}$ given by equations A54 and A55 in Supplementary File S1 yields:

$$\Delta_{\text{purge}}\overline{\sigma} \approx \frac{U^2}{\rho_{\text{h},\sigma z}}\left[3+\frac{7\,(Q-2)}{4}\right]\left(\frac{4U}{Q\overline{\varsigma}}\right)^{-\frac{2}{Q}}V_\sigma \tag{51}$$

where the term in $Q-2$ between brackets corresponds to the term on the second line of equation 44 (effects of additive-by-dominance and dominance-by-dominance epistasis).

Figure 5 shows simulation results obtained under Gaussian stabilizing selection acting on different numbers of traits n (the mean deleterious effect of mutations $\overline{\varsigma}$ being kept constant by adjusting the variance of mutational effects a^2). Under stabilizing

selection, inbreeding depression reaches an upper limit as the mutation rate U increases (this upper limit being lower for smaller values of n), explaining why high values of δ' could not be explored in Figure 5. Again, epistasis increases the parameter range under which selfing can invade (the effect of epistasis being stronger when the number of selected traits n is lower), and the QLE model yields correct predictions as long as inbreeding depression (and thus U) is not too large. In contrast with the previous example (uniformly deleterious alleles), the model predicts that $\Delta_{\text{purge}}\overline{\sigma}$ stays negligible, the difference between the dotted and solid/dashed curves in Figure 5 being mostly due to $\Delta_{\text{LD}}\overline{\sigma}$: selfers thus benefit from the fact that they can maintain beneficial combinations of alleles (mutations with compensatory effects) at different loci. Interestingly, for $n = 5$ and sufficiently high rates of pollen discounting κ, selfing can invade if inbreeding depression is lower than a given threshold, or is very high. The latter case corresponds to a situation where polymorphism is important (high U) and where large numbers of compensatory combinations of alleles are possible. Although the model predicts that the same phenomenon should occur for higher values of n, it was not observed in simulations with $n = 15$ and $n = 30$, except for $n = 15$ and $\kappa = 0.4$. However, Supplementary Figures S6 and S7 show that the evolution of selfing above a threshold value of δ' occurs more frequently when the fitness peak is flatter ($Q > 2$), and when mutations affecting the selfing rate have larger effects.

Finally, Figure 6 provides additional results on the effect of the number of selected traits n, for fixed values of the overall mutation rate U. Inbreeding depression is little affected by epistatic interactions when n is large, while low values of n tend to decrease inbreeding depression, explaining the shapes of the dotted curves showing the maximum level of pollen discounting for selfing to spread, when only taking into

account the effects of the automatic advantage and inbreeding depression. The difference between the dotted and solid/dashed curves shows the additional effect of linkage disequilibria generated by epistasis ($\Delta_{\mathrm{LD}}\overline{\sigma}$), whose relative importance increases as the number of traits n decreases, and as the mutation rate U increases. Because U stays moderate ($U = 0.2$ or 0.5), the analytical model provides accurate predictions of the parameter range over which selfing is favored.

DISCUSSION

The automatic transmission advantage associated with selfing and inbreeding depression are the two most commonly discussed genetic mechanisms affecting the evolution of self-fertilization. When these are the only forces at play, a selfing mutant arising in an outcrossing population is expected to increase in frequency as long as inbreeding depression is weaker than the automatic advantage, whose magnitude depends on the level of pollen discounting (Lande and Schemske, 1985; Holsinger et al., 1984). However, because selfers also tend to carry better purged genomes due to their increased homozygosity, several models showed that selfing mutants may invade under wider conditions than those predicted solely based on these two aforementioned forces (Charlesworth et al., 1990; Uyenoyama and Waller, 1991; Epinat and Lenormand, 2009; Porcher and Lande, 2005b; Gervais et al., 2014). Our analytical and simulation results confirm that the advantage procured through purging increases with the strength of selection against deleterious alleles and with the degree of linkage within the genome. The simulation results also indicate that the verbal prediction, according to which mutations causing complete selfing may invade a population independently of its level of

inbreeding depression (Lande and Schemske, 1985, p. 33), only holds when deleterious alleles have strong fitness effects, so that purging occurs rapidly (Figure 2D).

Whether purging efficiency should significantly contribute to the spread of selfing mutants depends on the genetic architecture of inbreeding depression. To date, experimental data point to a small contribution of strongly deleterious alleles to inbreeding depression: for example, Baldwin and Schoen (2019) recently showed that in the self-incompatible species *Leavenworthia alabamica*, inbreeding depression is not affected by three generations of enforced selfing (which should have lead to the elimination of deleterious alleles with strong fitness effects). Previous experiments on different plant species also indicate that inbreeding depression is probably generated mostly by weakly deleterious alleles (Dudash et al., 1997; Willis, 1999; Carr and Dudash, 2003; Charlesworth and Willis, 2009). Data on the additive variance in fitness within populations are also informative regarding the possible effect of purging: indeed, using our general expression for fitness (equation 8) and neglecting linkage disequilibria, one can show that the additive component of the variance in fitness in a randomly mating population (more precisely, the variance in $W/\overline{W}$) is given by the sum over selected loci of $2{a_j}^2 p_j q_j$ (see also eq. A3b in Charlesworth and Barton, 1996), a term which also appears in the effect of purging on the strength of selection for selfing (equation 36). Although estimates of the additive variance in fitness in wild populations remain scarce, the few estimates of the "evolvability" parameter (corresponding to the additive component of the variance in $W/\overline{W}$) available from plant species are small, of the order of a few percent (Hendry et al., 2018). Note that strictly, the effect of purging on the strength of selection for selfing is proportional to the quantity $\sum_j a_j \left[a_j + a_{j,j}\left(1-2p_j\right)\right] p_j q_j$ (equation 36), which may be larger than $\sum_j {a_j}^2 p_j q_j$

(for example, in the case of deleterious alleles with fixed s and h, the first quantity is approximately $s\,(1-h)\,U$ and the second shU). As explained in the Results section (and in Supplementary File S3), the strength of selection for selfing through purging may, in principle, be estimated from the increase in mean fitness following a single generation of selfing. However, the small values of the available estimates of $\sum_j {a_j}^2 p_j q_j$, together with the experimental evidence mentioned above on the genetics of inbreeding depression, indicate that selfing mutants probably do not benefit greatly from purging. Nevertheless, it remains possible that the strength of selection against deleterious alleles (a_j) increases in harsher environments (Cheptou et al., 2000; Agrawal and Whitlock, 2010), leading to stronger purging effects in such environments.

The effects of epistasis between deleterious alleles on inbreeding depression and on the evolution of mating systems have been little explored (but see Charlesworth et al., 1991). In this paper, we derived general expressions for the effect of epistasis between pairs of loci on inbreeding depression and on the strength of selection for selfing, that can be applied to more specific models. Our results show that different components of epistasis have different effects on inbreeding depression: negative dominance-by-dominance epistasis directly increases inbreeding depression due to covariances in homozygosity across loci among selfed offspring, while additive-by-additive and additive-by-dominance epistasis may indirectly affect inbreeding depression by changing the effective strength of selection (a_j) or effective dominance ($a_{j,j}$) of deleterious alleles. Very little is known on the average sign and relative magnitude of these different forms of epistasis. In principle, the overall sign of dominance-by-dominance effects can be deduced from the shape of the relation between the inbreeding coefficient of individuals (F) and their fitness (Crow and Kimura, 1970, p. 80), an accelerating

decline in fitness as F increases indicating negative e_{dxd}. The relation between F and fitness-related traits was measured in several plant species; the results often showed little departure from linearity (e.g., Willis, 1993; Kelly, 2005), but the experimental protocols used may have generated biases against finding negative e_{dxd} (Falconer and Mackay, 1996; Lynch and Walsh, 1998; Sharp and Agrawal, 2016).

Most empirical distributions of epistasis between pairs of mutations affecting fitness have been obtained from viruses, bacteria and unicellular eukaryotes (e.g., Martin et al., 2007; Kouyos et al., 2007; de Visser and Elena, 2007). While no clear conclusion emerges regarding the average coefficient of epistasis (some studies find that it is negative, other positive and other close to zero), a general observation is that epistasis is quite variable across pairs of loci. This variance of epistasis may slightly increase inbreeding depression when it remains small (by reducing the efficiency of selection against deleterious alleles, Phillips et al., 2000; Abu Awad and Roze, 2018), or decrease inbreeding depression when it is larger and/or effective recombination is sufficiently weak, so that selfing can maintain beneficial multilocus genotypes (Lande and Porcher, 2015; Abu Awad and Roze, 2018). Besides this "short-term" effect on inbreeding depression, the variance of epistasis also favors selfing through the progressive buildup of linkage disequilibria that increase mean fitness (associations between alleles with compensatory effects at different loci): this effect is equivalent to selection for reduced recombination rates caused by the variance of epistasis among loci, previously described by Otto and Feldman (1997). Interestingly, this effect may become stronger than inbreeding depression above a threshold value of the rate of mutation on traits under stabilizing selection (Figures 4, S7). Is the variance of epistasis typically large enough, so that the benefit of maintaining beneficial combinations of alleles may

significantly help selfing mutants to spread? Answering this question is difficult without better knowledge of the importance of epistatic interactions on fitness in natural environments. Nevertheless, some insights can be gained from our analytical results: for example, neglecting additive-by-dominance and dominance-by-dominance effects, equations 43 and D7 indicate that the effect of linkage disequilibria on the strength of selection for selfing should scale with the sum over pairs of selected loci of ${a_{jk}}^2 p_j q_j p_k q_k$, which also corresponds to the epistatic component of the variance in fitness in randomly mating populations. Although estimates of epistatic components of variance remain scarce, they are typically not larger than additive components (e.g., Hill et al., 2008), suggesting that the benefit of maintaining beneficial multilocus genotypes may be generally limited (given that the additive variance in fitness seems typically small, as discussed previously).

Previous models on the evolution of recombination showed that increased recombination rates may be favored when epistasis is negative and sufficiently weak relative to the strength of directional selection (e.g., Barton, 1995). Similarly, weakly negative epistasis may favour the maintenance of outcrossing (through selection for recombination): this effect does not appear in our model, because it involves higher order terms (proportional to $a_j a_k a_{jk}$) that were neglected in our analysis. When epistasis is weak ($a_{jk} \ll a_j,\ a_k$), these terms may become of the same order as the term in ${a_{jk}}^2$ arising in $\Delta_{\mathrm{LD}}\overline{\sigma}$, leading to a net effect of epistasis favoring outcrossing; however, the overall effect of epistasis should be negligible (relative to the effects of inbreeding depression and purging) when $a_{jk} \ll a_j,\ a_k$. Selection for recombination (and outcrossing) due to negative epistasis may become stronger when effective rates of recombination between loci become smaller (in particular, in highly selfing popula-

tions), which may prevent evolution towards complete selfing. Indeed, Charlesworth et al. (1991) showed that in the presence of negative epistasis between deleterious alleles, and when outcrossing is not stable, a selfing rate slightly below one corresponds to the evolutionarily stable strategy (ESS). In finite populations, selection for recombination is also driven by the Hill-Robertson effect (through a term proportional to ${a_j}^2{a_k}^2/N_\mathrm{e}$, e.g., Barton and Otto, 2005) even in the absence of epistasis. Again, while this term should generally stay negligible (relative to inbreeding depression and purging) in an outcrossing population, it may become more important in highly selfing populations, due to their reduced effective size. Accordingly, Kamran-Disfani and Agrawal (2014) showed that selfing rates slightly below one are selectively favoured over complete selfing in finite populations under multiplicative selection (no epistasis). Similar effects must have occurred in our simulations, although we did not check that selfing rates slightly below one resulted from selection to maintain low rates of outcrossing, rather than from the constant input of mutations at selfing modifier loci (this could be done by comparing the probabilities of fixation of alleles coding for different selfing rates, as in Kamran-Disfani and Agrawal, 2014). Together with the results of Charlesworth et al. (1991) and Kamran-Disfani and Agrawal (2014), our simulation results indicate that, while selection for recombination may favour the maintenance of low rates of outcrossing in highly selfing populations, it cannot explain the maintenance of mixed mating systems (involving higher rates of outcrossing) under constant environmental conditions (the selfing rate always evolved towards values either close to zero or one in our simulations). It is possible that mixed mating systems may be more easily maintained under changing environmental conditions, however (for example, under directional selection acting on quantitative traits); this represents an interesting avenue

for future research.

Acknowledgements: We thank the bioinformatics and computing service of Roscoff's Biological Station (Abims platform) for computing time, and Nick Barton, Sylvain Gandon and an anonymous reviewer for useful comments. This work was supported by the French Agence Nationale de la Recherche (project SEAD, ANR-13-ADAP-0011 and project SexChange, ANR-14-CE02-0001). Diala Abu Awad was partly funded by the TUM University Foundation Fellowship.

LITERATURE CITED

Abu Awad, D. and D. Roze. 2018. Effects of partial selfing on the equilibrium genetic variance, mutation load, and inbreeding depression under stabilizing selection. Evolution 72:751–769.

Agrawal, A. F. and M. C. Whitlock. 2010. Environmental duress and epistasis: how does stress affect the strength of selection on new mutations? Trends Ecol. Evol. 25:450–458.

Baldwin, S. J. and D. J. Schoen. 2019. Inbreeding depression is difficult to purge in self-incompatible populations of *Leavenworthia alabamica*. New Phytol. doi: 10.1111/nph.15963.

Barrett, S. C. H., R. Arunkumar, and S. I. Wright. 2014. The demography and population genomics of evolutionary transitions to self-fertilization in plants. Phil. Trans. Roy. Soc. (Lond.) B 369:20130344.

Barton, N. H. 1995. A general model for the evolution of recombination. Genet. Res. 65:123–144.

Barton, N. H. and S. P. Otto. 2005. Evolution of recombination due to random drift. Genetics 169:2353–2370.

Barton, N. H. and M. Turelli. 1991. Natural and sexual selection on many loci. Genetics 127:229–255.

Busch, J. W. and L. F. Delph. 2012. The relative importance of reproductive assurance

and automatic selection as hypotheses for the evolution of self-fertilization. Ann. Bot. 109:553–562.

Carr, D. E. and M. R. Dudash. 2003. Recent approaches into the genetic basis of inbreeding depression in plants. Phil. Trans. Roy. Soc. (Lond.) B 358:1071–1084.

Charlesworth, B. 1980. The cost of sex in relation to the mating system. J. Theor. Biol. 84:655–671.

Charlesworth, B. and N. H. Barton. 1996. Recombination load associated with selection for increased recombination. Genet. Res. 67:27–41.

Charlesworth, B., M. T. Morgan, and B. Charlesworth. 1990. Inbreeding depression, genetic load, and the evolution of outcrossing rates in a multilocus system with no linkage. Evolution 44:1469–1489.

Charlesworth, B., M. T. Morgan, and D. Charlesworth. 1991. Multilocus models of inbreeding depression with synergistic selection and partial self-fertilization. Genet. Res. 57:177–194.

Charlesworth, D. and B. Charlesworth. 1987. Inbreeding depression and its evolutionary consequences. Ann. Rev. Ecol. Syst. 18:237–268.

Charlesworth, D. and J. H. Willis. 2009. The genetics of inbreeding depression. Nat. Rev. Genet. 10:783–796.

Cheptou, P. O., E. Imbert, J. Lepart, and J. Escarre. 2000. Effects of competition on lifetime estimates of inbreeding depression in the outcrossing plant *Crepis sancta* (Asteraceae). J. Evol. Biol. 13:522–531.

Crow, J. F. and M. Kimura. 1970. An Introduction to Population Genetics Theory. Harper and Row, New York.

Darwin, C. 1876. The effects of cross- and self-fertilization in the vegetable kingdom. John Murray, London.

de Visser, J. A. G. M. and S. F. Elena. 2007. The evolution of sex: empirical insights into the roles of epistasis and drift. Nat. Rev. Genet. 8:139–149.

Dudash, M. R., D. E. Carr, and C. B. Fenster. 1997. Five generations of enforced selfing and outcrossing in *Mimulus guttatus*: inbreeding depression variation at the population and family level. Evolution 51:54–65.

Epinat, G. and T. Lenormand. 2009. The evolution of assortative mating and selfing with in- and outbreeding depression. Evolution 63:2047–2060.

Falconer, D. S. and T. F. C. Mackay. 1996. Introduction to Quantitative Genetics. Addison Wesley Longman, Harlow.

Fisher, R. 1941. Average excess and average effect of a gene substitution. Ann. Eugen. 11:53–63.

Gervais, C., D. Abu Awad, D. Roze, V. Castric, and S. Billiard. 2014. Genetic architecture of inbreeding depression and the maintenance of gametophytic self-incompatibility. Evolution 68:3317–3324.

Gros, P.-A., H. Le Nagard, and O. Tenaillon. 2009. The evolution of epistasis and its links with genetic robustness, complexity and drift in a phenotypic model of adaptation. Genetics 182:277–293.

Hendry, A. P., D. J. Schoen, M. E. Wolak, and J. M. Reid. 2018. The contemporary evolution of fitness. Ann. Rev. Ecol. Evol. Syst. 49:457–476.

Hill, W. G., M. E. Goddard, and P. M. Visscher. 2008. Data and theory point to mainly additive genetic variance for complex traits. PLoS Genetics 4:e1000008.

Holsinger, K. E. 1991. Mass-action models of plant mating systems: the evolutionary stability of mixed mating systems. Am. Nat. 138:606–622.

Holsinger, K. E., M. W. Feldman, and F. B. Christiansen. 1984. The evolution of self-fertilization in plants: a population genetic model. Am. Nat. 124:446–453.

Igic, B. and J. W. Busch. 2013. Is self-fertilization an evolutionary dead end? New Phytol. 198:386–397.

Jarne, P. and J. R. Auld. 2006. Animals mix it up too: the distribution of self-fertilization among hermaphroditic animals. Evolution 60:1816–1824.

Kamran-Disfani, A. and A. F. Agrawal. 2014. Selfing, adaptation and background selection in finite populations. J. Evol. Biol. 27:1360–1371.

Kelly, J. K. 2005. Epistasis in monkeyflowers. Genetics 171:1917–1931.

Kirkpatrick, M., T. Johnson, and N. H. Barton. 2002. General models of multilocus evolution. Genetics 161:1727–1750.

Kouyos, R. D., O. K. Silander, and S. Bonhoeffer. 2007. Epistasis between deleterious mutations and the evolution of recombination. Trends Ecol. Evol. 22:308–315.

Lande, R. and E. Porcher. 2015. Maintenance of quantitative genetic variance under partial self-fertilization, with implications for the evolution of selfing. Genetics 200:891–906.

Lande, R. and D. W. Schemske. 1985. The evolution of self-fertilization and inbreeding depression in plants. I. Genetic models. Evolution 39:24–40.

Lynch, M. and J. B. Walsh. 1998. Genetics and Analysis of Quantitative Traits. Sinauer Associates, Sunderland, MA.

Manna, F., G. Martin, and T. Lenormand. 2011. Fitness landscapes: an alternative theory for the dominance of mutation. Genetics 189:923–937.

Martin, G., S. F. Elena, and T. Lenormand. 2007. Distributions of epistasis in microbes fit predictions from a fitness landscape model. Nat. Genet. 39:555–560.

Martin, G. and T. Lenormand. 2006a. The fitness effect of mutations across environments: a survey in light of fitness landscape models. Evolution 60:2413–2427.

———. 2006b. A general multivariate extension of Fisher's geometrical model and the distribution of mutation fitness effects across species. Evolution 60:893–907.

Nagylaki, T. 1976. A model for the evolution of self-fertilization and vegetative reproduction. J. Theor. Biol. 58:55–58.

———. 1993. The evolution of multilocus systems under weak selection. Genetics 134:627–647.

Nordborg, M. 1997. Structured coalescent processes on different time scales. Genetics 146:1501–1514.

Otto, S. P. and M. W. Feldman. 1997. Deleterious mutations, variable epistatic interactions, and the evolution of recombination. Theor. Popul. Biol. 51:134–47.

Phillips, P. C., S. P. Otto, and M. C. Whitlock. 2000. Beyond the average: the evolutionary importance of gene interactions and variability of epistatic effects. Pp. 20–38 *in* J. B. Wolf, E. D. Brodie, and M. J. Wade, eds. Epistasis and the Evolutionary Process. Oxford University Press, New York.

Porcher, E. and R. Lande. 2005a. The evolution of self-fertilization and inbreeding depression under pollen discounting and pollen limitation. J. Evol. Biol. 18:497–508.

———. 2005b. Loss of gametophytic self-incompatibility with evolution of inbreeding depression. Evolution 59:46–60.

Roze, D. 2015. Effects of interference between selected loci on the mutation load, inbreeding depression and heterosis. Genetics 201:745–757.

Roze, D. and A. Blanckaert. 2014. Epistasis, pleiotropy and the mutation load in sexual and asexual populations. Evolution 68:137–149.

Sharp, N. P. and A. F. Agrawal. 2016. The decline in fitness with inbreeding: evidence for negative dominance-by-dominance epistasis in *Drosophila melanogaster*. J. Evol. Biol. 29:857–864.

Stebbins, G. L. 1957. Self fertilization and population variability in higher plants. Am. Nat. 91:337–354.

Stone, J. L., E. J. Van Wyk, and J. R. Hale. 2014. Transmission advantage favors

selfing allele in experimental populations of self-incompatible Witheringia solanacea (Solanaceae). Evolution 68:1845–1855.

Tenaillon, O., O. K. Silander, J.-P. Uzan, and L. Chao. 2007. Quantifying organismal complexity using a population genetic approach. PLoS One 2:e217.

Uyenoyama, M. K. and D. M. Waller. 1991. Coevolution of self-fertilization and inbreeding depression. I. Mutation-selection balance at one and two loci. Theor. Popul. Biol. 40:14–46.

Weir, B. S. and C. C. Cockerham. 1973. Mixed self and random mating at two loci. Genet. Res. 21:247–262.

Willis, J. H. 1993. Effects of different levels of inbreeding on fitness components in *Mimulus guttatus*. Evolution 47:864–876.

———. 1999. The role of genes of large effect on inbreeding depression in *Mimulus guttatus*. Evolution 53:1678–1691.

Table 1: Parameters and variables of the model.

σ	Selfing rate
$\overline{\sigma}$, V_σ	Mean and variance in the selfing rate in the population
κ	Rate of pollen discounting
ℓ_σ	Number of loci affecting the selfing rate
W, $\overline{W}$	Fitness of an individual, and average fitness
ℓ	Number of loci affecting fitness
U	Overall (haploid) mutation rate at loci affecting fitness
p_j, q_j	Frequencies of alleles 1 and 0 at loci affecting fitness
ℓ	Number of loci affecting selected traits
n_d	Mean number of deleterious alleles per haploid genome
s, h	Selection and dominance coefficients of deleterious alleles
e_axa, e_axd, e_dxd	Additive-by-additive, additive-by-dominance and dominance-by-dominance epistasis between deleterious alleles
β	Strength of synergistic epistasis in Charlesworth et al.'s (1991) model
n	Number of quantitative traits under stabilizing selection
V_s	Strength of stabilizing selection
$r_{\alpha j}$	Effect of allele 1 at locus j on trait α
ς_j	Fitness effect of a heterozygous mutation in an optimal genotype (stabilizing selection model)

$\bar{\varsigma}$	Average fitness effect of a heterozygous mutation in an optimal genotype (stabilizing selection model)
a^2	Variance of mutational effects on traits under stabilizing selection
Q	Shape of the fitness peak (equation 15)
$a_{\mathbb{U},\mathbb{V}}$	Effect of selection on the sets $\mathbb{U}$ and $\mathbb{V}$ of loci present on the maternally and paternally inherited haplotypes of an individual (equation 8)
$D_{\mathbb{U},\mathbb{V}}$	Genetic association between the sets $\mathbb{U}$ and $\mathbb{V}$ of loci present on the maternally and paternally inherited haplotypes of an individual (equation 4)
ρ_{jk}	Recombination rate between loci j and k
U_{self}	Mutation rate at loci affecting the selfing rate
σ^2_{self}	Variance of mutational effects at loci affecting the selfing rate
δ	Inbreeding depression
δ'	Inbreeding depression measured after selection
F	Inbreeding coefficient
G_{jk}	Identity disequilibrium between loci j and k
G	Identity disequilibrium between freely recombining loci

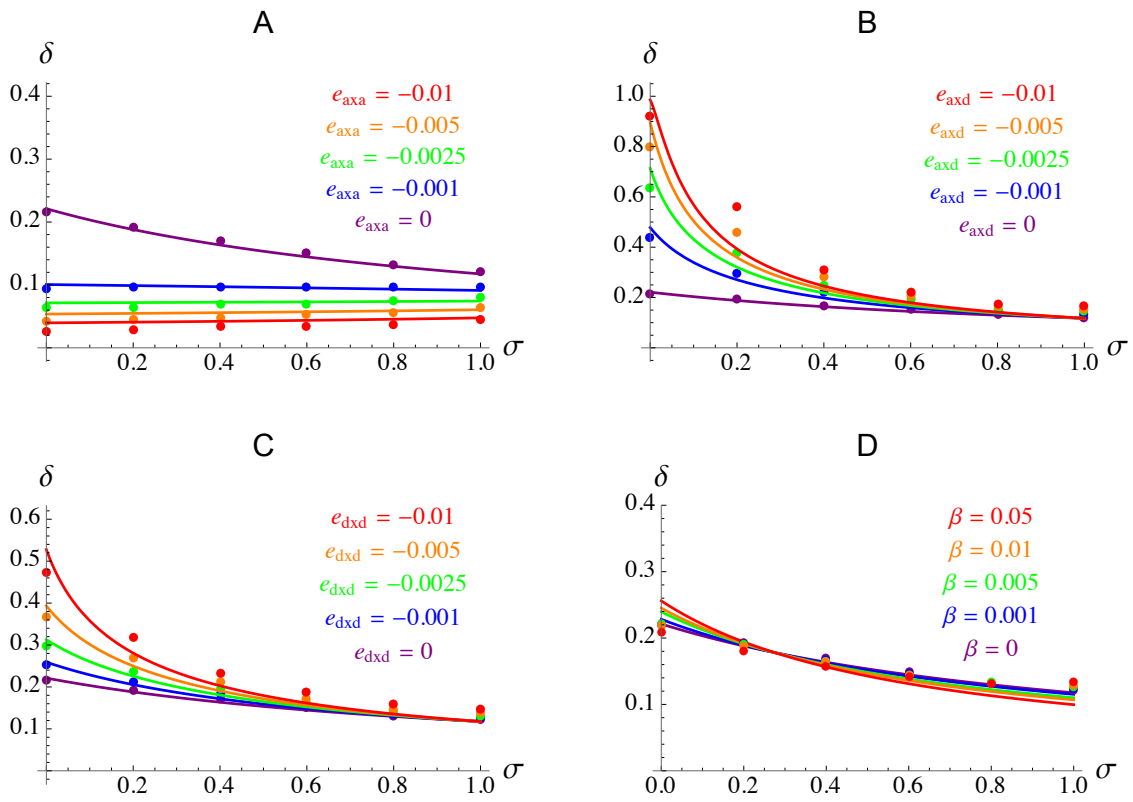

Figure 1. Inbreeding depression δ as a function of the selfing rate σ. A–C: effects of the different components of epistasis between deleterious alleles, additive-by-additive (e_{axa}), additive-by-dominance (e_{axd}) and dominance-by-dominance (e_{dxd}) — in each plot, the other two components of epistasis are set to zero. D: results obtained using Charlesworth et al.'s (1991) fitness function, where β represents synergistic epistasis between deleterious alleles (slightly modified as explained in Supplementary File S1). Dots correspond to simulation results (error bars are smaller than the size of symbols), and curves to analytical predictions from equations 24 and 27. Parameter values: $U = 0.25$, $s = 0.05$, $h = 0.25$. In the simulations $N = 20{,}000$ (population size) and $R = 10$ (genome map length); simulations lasted 10^5 generations and inbreeding depression was averaged over the last 5×10^4 generations.

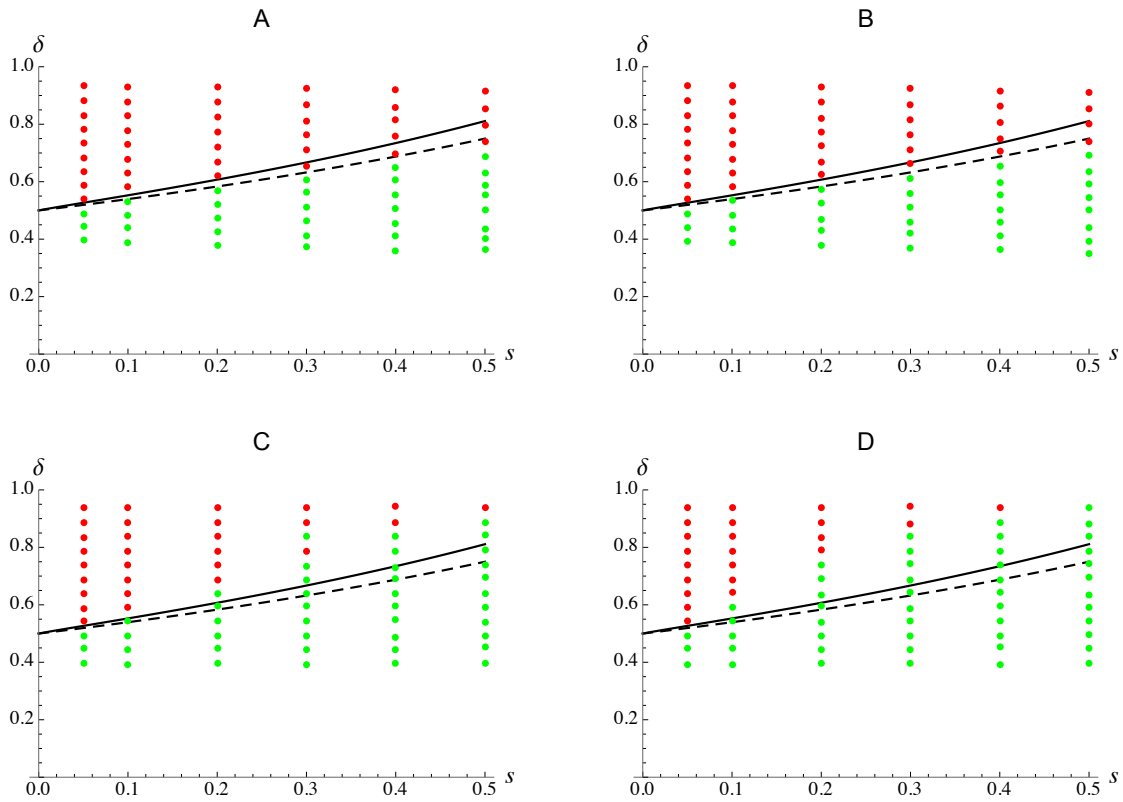

Figure 2. Evolution of selfing in the absence of epistasis. The solid curve shows the maximum value of inbreeding depression δ for selfing to spread in an initially outcrossing population, as a function of the strength of selection s against deleterious alleles (obtained from equations 35 and 37, after integrating equation 37 over the genetic map), while the dashed curve corresponds to the same prediction in the case of unlinked loci (obtained by setting $\rho_{ij} = 1/2$ in equation 37). Dots correspond to simulation results (using different values of U for each value of s, in order to generate a range of values of δ). In the simulations the population evolves under random mating during the first 20,000 generations (inbreeding depression is estimated by averaging over the last 10,000 generations); mutation is then introduced at the selfing modifier locus. A red dot means that the selfing rate stayed below 0.05 during the 2×10^5

generations of the simulation, while a green dot means that selfing increased (in which case the population always evolved towards nearly complete selfing). Parameter values: $\kappa = 0$, $h = 0.25$, $R = 10$; in the simulations $N = 20{,}000$, $U_{\text{self}} = 0.001$ (mutation rate at the selfing modifier locus). In A, the standard deviation of mutational effects at the modifier locus is set to $\sigma_{\text{self}} = 0.01$, while it is set to $\sigma_{\text{self}} = 0.03$ in B, and to $\sigma_{\text{self}} = 0.3$ in C. In D, only two alleles are possible at the modifier locus, coding for $\sigma = 0$ or 1, respectively.

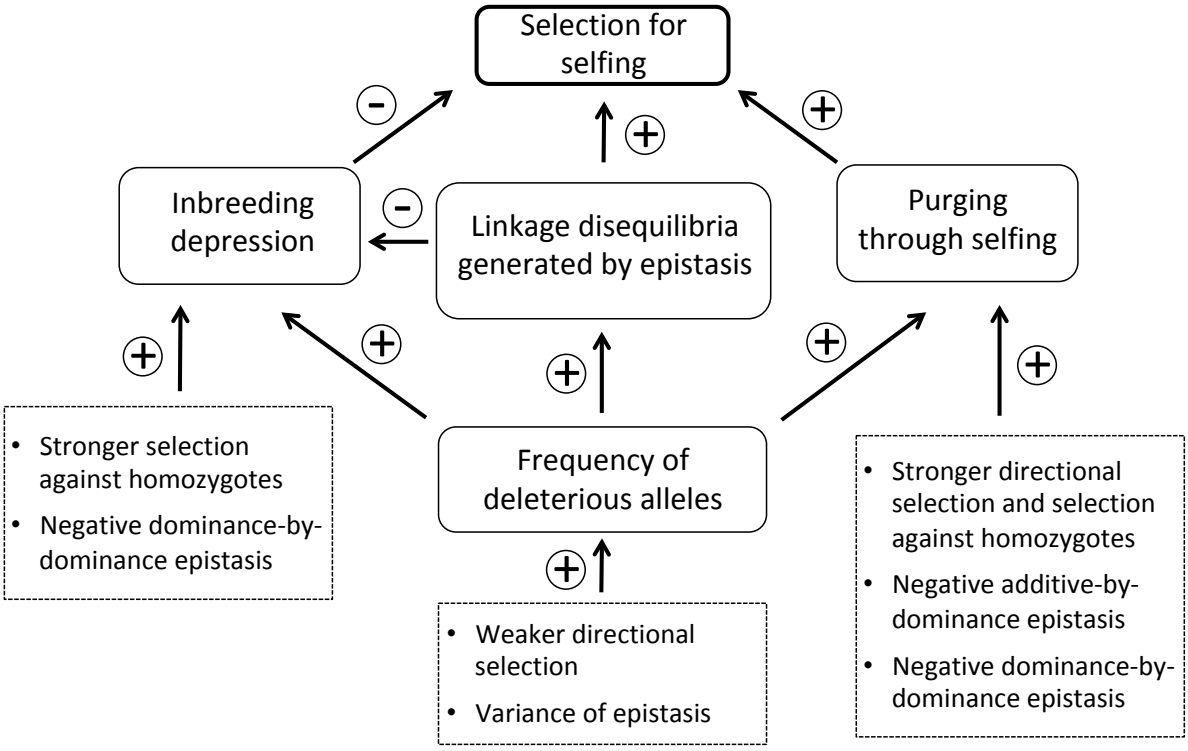

Figure 3. Summary of the effects of the strength of directional selection (a_j), the strength of selection against homozygotes ($a_{j,j}$) and epistasis (a_{jk}, $a_{j,k}$, $a_{jk,j}$, $a_{jk,jk}$) on the three components of indirect selection for selfing, in a randomly mating population. Note that a_j and $a_{j,j}$ may be affected by epistatic interactions among loci (e.g., equations A10, A11 in Supplementary File S1).

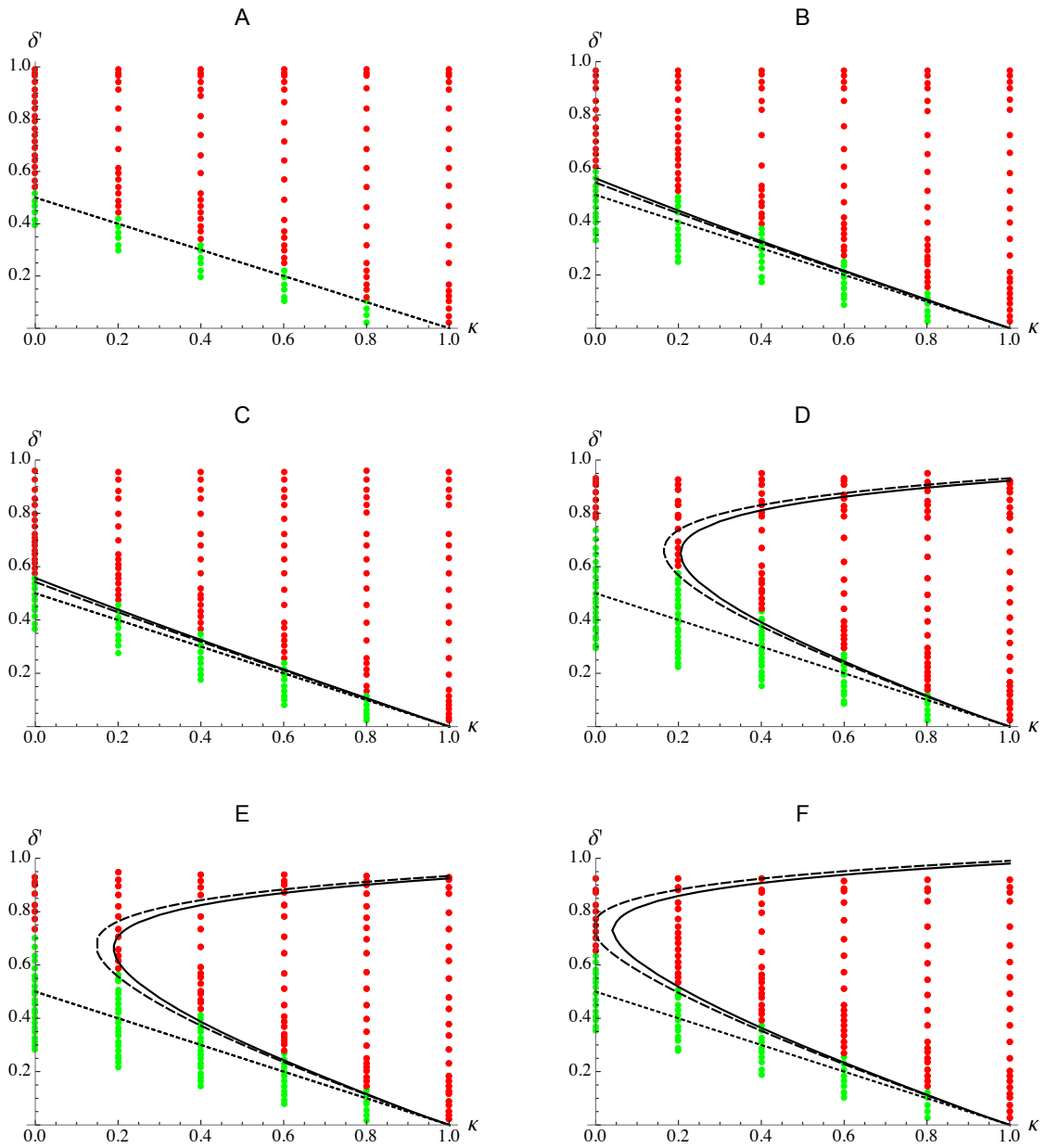

Figure 4. Evolution of selfing with fixed, negative epistasis. The different plots show the maximum value of inbreeding depression δ' (measured after selection) for selfing to spread in an initially outcrossing population, as a function of the rate of pollen discounting κ. Green and red dots correspond to simulation results and have the same meaning as in Figure 2 (δ' was estimated by averaging over the last 10,000 generations of the 20,000 preliminary generations without selfing, simulations lasted

2×10^5 generations). The dotted lines correspond to the predicted maximum inbreeding depression for selfing to increase obtained when neglecting $\Delta_{\mathrm{LD}}\overline{\sigma}$ and $\Delta_{\mathrm{purge}}\overline{\sigma}$ (that is, $\delta' = (1-\kappa)/2$), the dashed curves correspond to the prediction obtained using the expressions for $\Delta_{\mathrm{LD}}\overline{\sigma}$ and $\Delta_{\mathrm{purge}}\overline{\sigma}$ under free recombination (equations 46 and 49), while the solid curves correspond to the predictions obtained by integrating equations 45 and 48 over the genetic map (the effect of $\Delta_{\mathrm{LD}}\overline{\sigma}$ is predicted to be negligible relative to the effect of $\Delta_{\mathrm{purge}}\overline{\sigma}$ in all cases). To obtain these predictions, the relation between the mean number of deleterious alleles per haplotype n_{d} (that appears in equations 45–46 and 48–49) and δ' was obtained from a fit of the simulation results. A: $e_{\mathrm{axa}} = e_{\mathrm{axd}} = e_{\mathrm{dxd}} = 0$; B: $e_{\mathrm{axa}} = e_{\mathrm{dxd}} = 0$, $e_{\mathrm{axd}} = -0.01$; C: $e_{\mathrm{axa}} = e_{\mathrm{axd}} = 0$, $e_{\mathrm{dxd}} = -0.01$; D: $e_{\mathrm{axa}} = -0.005$, $e_{\mathrm{axd}} = -0.01$, $e_{\mathrm{dxd}} = 0$; E: $e_{\mathrm{axa}} = -0.005$, $e_{\mathrm{axd}} = e_{\mathrm{dxd}} = -0.01$; F: Charlesworth et al.'s (1991) model with $\beta = 0.05$. Other parameter values: $s = 0.05$, $h = 0.25$, $R = 20$; in the simulations $N = 20{,}000$, $U_{\mathrm{self}} = 0.001$ (mutation rate at the selfing modifier locus), $\sigma_{\mathrm{self}} = 0.03$ (standard deviation of mutational effects at the modifier locus).

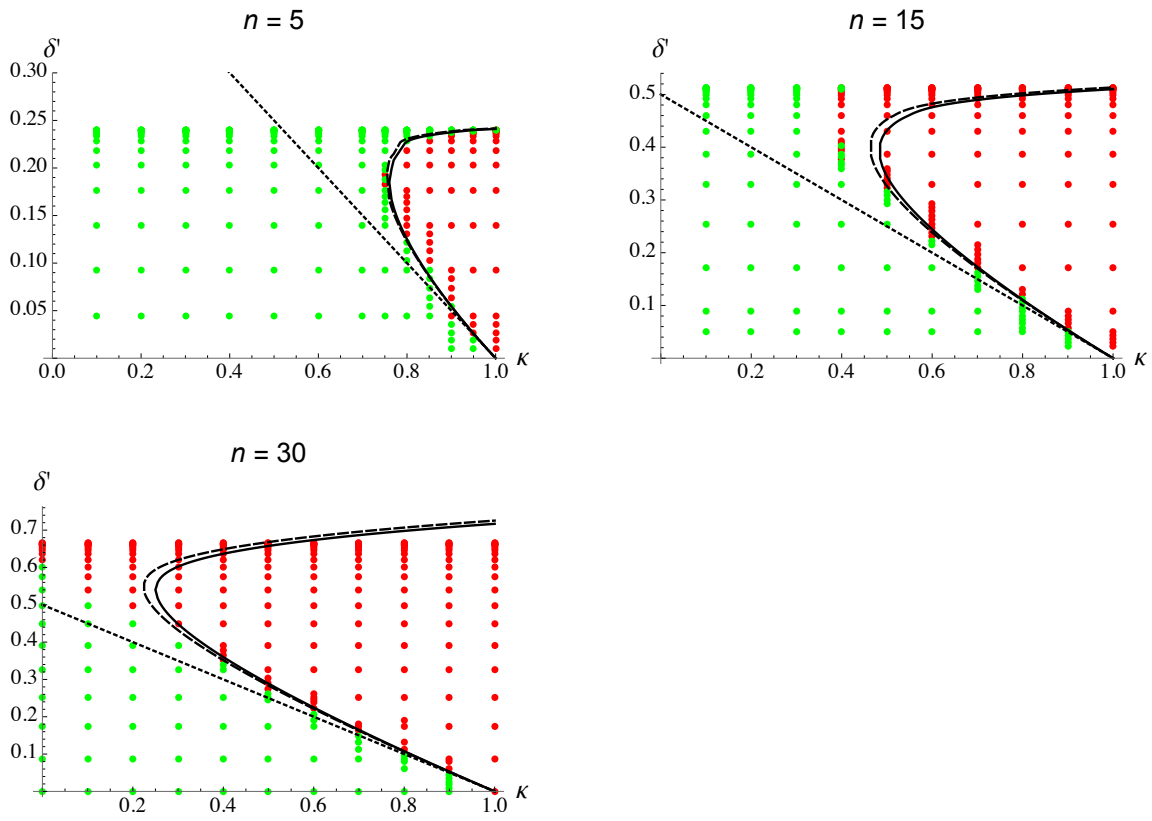

Figure 5. Evolution of self-fertilization under Gaussian stabilizing selection. The three plots show the effects of inbreeding depression δ' (measured after selection) and pollen discounting (parameter κ) on the evolution of self-fertilization, for different numbers of traits under selection ($n = 5$, 15 and 30). Green and red dots correspond to simulation results and have the same meaning as in Figures 2 and 4 (δ' was estimated by averaging over the last 10,000 generations of the 20,000 preliminary generations without selfing, simulations lasted 5×10^4 generations). The fact that inbreeding depression reaches a plateau as U increases (at lower values of δ' for lower values of n) sets an upper limit to the values of δ' that can be obtained in the simulations. The dotted lines correspond to the predicted maximum inbreeding depression for selfing to increase obtained when neglecting $\Delta_{\mathrm{LD}}\overline{\sigma}$ and $\Delta_{\mathrm{purge}}\overline{\sigma}$ (that is, $\delta' = (1-\kappa)/2$), the

dashed curves correspond to the prediction obtained using the expression for $\Delta_{\mathrm{LD}}\overline{\sigma}$ under free recombination (that is, $6U^2V_\sigma/n$, see equation 50), while the solid curves correspond to the predictions obtained by integrating equation 50 over the genetic map (the effect of $\Delta_{\mathrm{purge}}\overline{\sigma}$ is predicted to be negligible relative to the effect of $\Delta_{\mathrm{LD}}\overline{\sigma}$). To obtain these predictions, the relation between U and δ' was obtained from a fit of the simulation results. Other parameter values: $\overline{\varsigma} = 0.01$, $R = 20$; in the simulations $N = 5{,}000$, $U_{\mathrm{self}} = 0.001$ (overall mutation rate at selfing modifier loci), $\sigma_{\mathrm{self}} = 0.01$ (standard deviation of mutational effects on selfing).

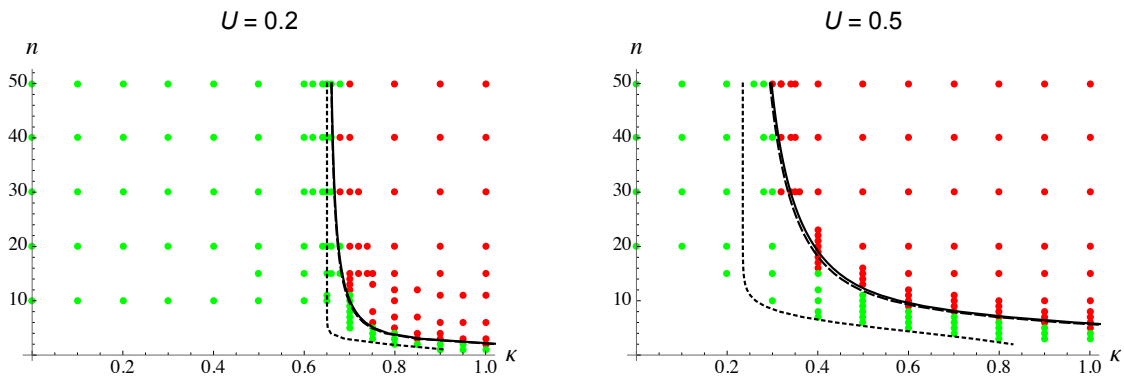

Figure 6. Evolution of self-fertilization under Gaussian stabilizing selection. The two plots show the effect of the number of traits under selection n and pollen discounting (parameter κ) on the evolution of self-fertilization for two values of the mutation rate on traits under stabilizing selection ($U = 0.2$ and 0.5). Green and red dots correspond to simulation results and have the same meaning as in the previous figures. The dotted curves show the maximum value of pollen discounting κ for selfing to increase obtained when neglecting $\Delta_{\mathrm{LD}}\overline{\sigma}$ and $\Delta_{\mathrm{purge}}\overline{\sigma}$ (that is, $\delta' = (1-\kappa)/2$), while the dashed and solid curves correspond to the predictions including the term $\Delta_{\mathrm{LD}}\overline{\sigma}$ (from equation 50) under free recombination (dashed) or integrated over the genetic map (solid). To obtain these predictions, the relation between n and δ' was obtained from a fit of the simulation results. Other parameter values: $\overline{\varsigma} = 0.01$, $R = 20$; in the simulations $N = 5{,}000$, $U_{\mathrm{self}} = 0.001$ (overall mutation rate at selfing modifier loci), $\sigma_{\mathrm{self}} = 0.01$ (standard deviation of mutational effects on selfing).